Modelling the Relationship between Parental Behaviour and Childhood Skill Development:
Empirical Evidence from the UK and Canada

Ashton Mary Pittendreigh Brown

Fitzwilliam College

Faculty of Education
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

September 2019

Declaration

- This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.
- It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.
- I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.
- It does not exceed the prescribed word limit for the relevant Degree Committee.

Ashton Mary Pittendreigh Brown
September 2019

Acknowledgements

My time as a doctoral student has been filled with many adventures – and a few misadventures. Throughout it all, I have been incredibly fortunate to receive support, guidance and friendship from countless individuals.

I am especially grateful to my supervisor, Professor Anna Vignoles who encouraged me to embark on this project and has guided me throughout the process. I have been inspired by her incredible intellect and her passion for research. Anna's insight and feedback has made me a better researcher. I thank her for being patient with me throughout the PhD; challenging me to demand more from myself; and supporting me when I didn't believe in my own abilities.

The Canadian portion of this thesis would never have been possible without the support of Dr. Richard Wanner and the Prarie Research Data Centre at the University of Calgary. Many thanks to Jeanne Williams, Dina Lavorato, and Charlie Victorino for their friendship and assistance during the four months I spent analysing my Canadian data. I am especially grateful to Jeanne for her encouragement and kindness, not to mention her invaluable assistance as my 'local-analyst'.

I am deeply indebted to several organisations that have played central roles in my Cambridge experience. Thanks to the *Cambridge Trust* and the *Canada-UK Foundation* for funding my research and tuition fees; and to *Fitzwilliam College* for not only supporting me financially through their PhD Studentship, but also cheering me on in all my endeavours and becoming my home in Cambridge. Likewise, the *Cambridge University Women's Boat Club* helped define my time as a Cambridge student. Thank you for driving me to find an inner strength that I didn't know I possessed; and for reminding me of the importance of true friendship.

I must also thank the many people who have given advice, read drafts, listened to my research ideas and guided me throughout the process. Special thanks to Jennifer, Laura, and Sarah for being a phenomenal team of proof-readers; and to Bjoern, Gill and Tanya for keeping me company throughout the process of writing up.

Beyond those who helped with the thesis itself, there are many other people who have shaped my time in Cambridge. While I cannot name everyone, I owe a debt of gratitude to all those who helped me make my way through this journey. I would especially like to thank my 'Cambridge Family' — Ciara, Paul, Laura and Conor — you fell into my life at just the right time and I am forever grateful for your love and support.

Finally, I need to thank my family, for their unconditional love throughout all my adventures. Mom and Dad, thank you for raising me to have a passion for learning and instilling the importance of kindness and social justice that guides my research. Thanks to my siblings, Fraser and Emily for putting up with my status as absentee sister, sending me adorable 'pup-dates' when I need a smile and supporting me in so many ways.

Abstract

Thesis Title: Modelling the Relationship between Parental Behaviour and Childhood Skill Development: Empirical Evidence from the UK and Canada

Author: Ashton Mary Pittendreigh Brown

This longitudinal, quantitative methods research defines a theoretical framework to model the developmental trajectories of cognitive and non-cognitive skills in primary school children and capture the role that parenting plays in the joint evolution of these skills. This framework draws on theoretical models and empirical findings from the fields of economics, psychology and education. Specifically, this thesis adapts the economic framework of Cunha and Heckman (2007, 2008) to separately measure the effect of financial resources and of inputs to development in the form of parenting behaviours; to allow for literature from psychology and education to aid in the identification of multiple types of such parental inputs; and to examine the differing impacts from each of these types of investment. Distinguishing between financial investments in children and other parental inputs to development is critical for the creation of public policies which address existing socio-economic disparities in child development.

Applying this proposed framework to nationally representative, longitudinal survey data from the UK and Canada, yields empirical estimates of skill formation in two contexts. These studies demonstrate how the model can be adapted to examine various aspects of parenting or to accommodate different types of longitudinal measures. Additionally, each empirical application provides precise estimates of how cognitive and non-cognitive skills evolve from birth to adolescence; and specific measurements of how the time that parents spend with their children impacts skill development.

In both the UK and Canadian data, three parenting constructs are identified. Each of these inputs has a significant effect on the development of both cognitive and non-cognitive skills, with differing periods of sensitivity to each type of parental input. Unlike past research which finds that as children age parental input becomes ineffective in promoting cognitive development, the results presented in this thesis find that some types of parental input are most effective for cognitive development in early childhood while other types have measurable effects on cognitive ability in older children.

Keywords: parenting; skill development; cognitive skills; non-cognitive skills; developmental trajectories; quantitative

Table of contents

List of figures **xiii**

List of tables **xv**

List of abbreviations **xvii**

1 Introduction **1**
1.1 Research Aims . . . 3
1.2 Overview of the Thesis . . . 4
Organisation and Structure . . . 4

2 Literature Review **7**
2.1 Clarifying Relevant Terminology . . . 8
Defining Skills and Abilities . . . 8
Cognitive Ability . . . 9
Non-Cognitive Ability . . . 9
Defining Parental Input and Investment: . . . 10
2.2 Existing Empirical Research . . . 11
The Predictive Power of Early Skills (Why do Skills Matter?) . . . 11
Empirical Studies of Skill Development . . . 12
2.3 Theoretical Concepts: Economics . . . 16
Using Human Capital Theory to Model Development . . . 16
Applying Production Functions to Skill Formation . . . 18
Early Research Using Production Functions . . . 18
Recent Research Using Production Functions . . . 20
Key Considerations for Contemporary Production Functions . . . 20
Other Theoretical Considerations from Economics . . . 24
Socio-Economic Disadvantage and Skill Development . . . 24
Separating Parental Motivations from Parental Investment . . . 27
Theoretical Implications of Formal Education . . . 28
2.4 Theoretical Concepts: Psychology . . . 30
Defining Skills Using Psychological Theories . . . 30
Theoretical Models of Child Development . . . 37
Theoretical Models of Parenting . . . 41
2.5 Using Theory: Education Research . . . 46
School Readiness . . . 47

2.6 Implications for this Thesis . . . 50
Relevant Considerations from Economics . . . 50
Relevant Considerations from Psychology . . . 50
Relevant Considerations from Education . . . 51
Consolidating these Fields . . . 52

3 Theoretical Framework 53
3.1 Defining the Conceptual Model . . . 54
Modelling Skill Development Using a Production Function . . . 55
Defining Human Capital in Children . . . 55
Defining Parental Input to Child Development . . . 56
Choosing the Correct Specification for the Production Function . . . 57
3.2 Full Model Specification . . . 59
Structural Model . . . 59
Measurement Model . . . 66
Combining the Structural and Measurement Models . . . 71
Conditions Necessary for Identification of Full Model . . . 72
3.3 Defining Parental Investment . . . 74
3.3.1 Indicator Selection: Relevant Measures of Parental Input . . . 74
3.3.2 Dimensionality: Setting the Number of Factors . . . 77
Existing Approaches: . . . 77
Using Statistical Analysis to Determine Dimensionality . . . 78
Factor Analysis versus Data Reduction . . . 78
Using EFA to Determine Number of Factors: . . . 80
Interpreting EFA Results . . . 85
Ensuring that Estimates of Dimensionality are Conceptually Sound 86
3.3.3 Assigning the Factor Structure . . . 88
Accepted Best Practices for Identifying a Factor Solution . . . 88
Additional Considerations for this Thesis . . . 89
Resulting Factor Solution – Parental Investment . . . 89
3.4 Empirically Estimating the Model . . . 90
Data Requirements . . . 90
Sample Size . . . 90
Longitudinal Coverage and Relevant Data . . . 91
Statistical Considerations when using Data from Cohort Studies . . . 92
Understanding Sampling Strategies . . . 92
Item Non-Response and Attrition . . . 93
Adjusting for Sample Design, Item Non-Response and Attrition . . . 94
Analytical Procedures . . . 95
3.5 Chapter Summary . . . 96

4 Empirical Application I: United Kingdom 97
4.1 Introduction . . . 98
4.2 Data . . . 101
Data Selection . . . 101
Millennium Cohort Study (MCS) . . . 102
Survey Description . . . 102
Data Access . . . 103

Ethical Considerations . . . 103
Sample Size and Attrition . . . 104
Sample Design . . . 105
Data Collection . . . 106
Selection of Relevant Subsample . . . 107
Measures used in Present Analysis . . . 108
Demographic Measures . . . 108
Measures of Cognitive Ability . . . 114
Measures of Non-Cognitive Ability . . . 121
Measures of Parental Investment . . . 126
4.3 Empirical Model . . . 131
Structural Model . . . 131
Measurement Model . . . 132
Cognitive Skills . . . 132
Non-Cognitive Skills . . . 133
Parental Investment . . . 134
Analysis Procedures . . . 134
4.4 Results . . . 135
Measurement Model: Cognitive Ability . . . 135
Measurement Model: Non-Cognitive Ability . . . 137
Confirming the Factor Structure . . . 137
Results from the Measurement Model: . . . 137
Measurement Model: Parental Investment . . . 140
Determining the Factor Structure . . . 140
Results from the Measurement Model . . . 142
Structural Model: . . . 144
4.5 Discussion . . . 150
Support for Methodological Approach . . . 150
Empirical Evidence . . . 151

5 Empirical Application II: Canada **153**
5.1 Introduction . . . 154
5.2 Data . . . 157
Data Selection . . . 157
National Longitudinal Survey of Children and Youth (NLSCY) . . . 158
Survey Description . . . 158
Data Access . . . 159
Ethical Considerations . . . 160
Survey Sample . . . 161
Sample Design . . . 163
Data Collection . . . 165
Selection of Relevant Subsample . . . 165
Measures Used in Present Analysis . . . 166
Demographic Measures . . . 166
Cognitive Measures . . . 171
Non-Cognitive Measures . . . 176
Investment Measures . . . 183
5.3 Empirical Model . . . 188

Structural Model . . . 188
Measurement Model . . . 189
Cognitive Skills . . . 189
Non-Cognitive Skills . . . 189
Parental Investment . . . 190
Analysis Procedures . . . 192
5.4 Results . . . 193
Measurement Model: Cognitive Ability . . . 193
Measurement Model: Non-Cognitive Ability . . . 195
Confirming the Factor Structure . . . 195
Results from the Measurement Model . . . 195
Measurement Model: Parental Investment . . . 198
Determining the Factor Structure . . . 198
Results from the Measurement Model . . . 200
Structural Model: . . . 202
5.5 Discussion . . . 207

6 Discussion **209**
6.1 Summary of Findings . . . 210
General Overview . . . 210
Revisiting the Research Questions . . . 211
6.2 Contributions to the Field . . . 213
Methodological Contributions . . . 213
Empirical Contributions . . . 214
6.3 Implications for Policy and Practice . . . 215
6.4 Strengths and Limitations . . . 217
Methodological Strengths and Limitations . . . 217
Strengths and Limitations of the Data . . . 217
6.5 Directions for Future Research . . . 218
Future Applications of my Methodological Approach . . . 218
Building on my Empirical Findings . . . 219

References **221**

Appendix A Additional Details about MCS Data and Analysis **243**
A.1 Summary of MCS Survey Elements . . . 243
A.2 MCS Sampling Strategies . . . 251
A.3 Full Code for Analysis . . . 253

Appendix B Additional Information about the NLSCY **259**
B.1 Summary of NLSCY Survey Elements . . . 259
B.2 Application for Access to Canadian Data . . . 261
B.3 NLSCY Sampling Strategies . . . 273
B.4 Survey Weights . . . 275
B.5 Bootstrap Weights . . . 277
B.6 Full Code for Analysis . . . 279
B.7 Results for 5 Cycle Model . . . 284

List of figures

2.1 Contextual Model of Parenting . . . 42
2.2 Parenting Style Typologies . . . 43

3.1 Structural Model of Skill Formation with Two Investment Types . . . 63
3.2 Example of a Measurement Model for Cognitive Ability . . . 69
3.3 Full Dynamic Model of Skill Formation with Two Investment Types . . . 71
3.4 Comparing EFA and PCA . . . 79
3.5 Eigenvalue-Based Factor Retention Methods . . . 82

4.1 MCS at a Glance . . . 102
4.2 Distribution of Cognitive Scores: MCS Sweep 2 (Age 3) . . . 119
4.3 Distribution of Cognitive Scores: MCS Sweep 3 (Age 5) . . . 119
4.4 Distribution of Cognitive Scores: MCS Sweep 4 (Age 7) . . . 120
4.5 Distribution of Cognitive Scores: MCS Sweep 5 (Age 11) . . . 120
4.6 Distribution of SDQ Scores: MCS Sweep 2 (Age 3) . . . 125
4.7 Distribution of SDQ Scores: MCS Sweep 3 (Age 5) . . . 125
4.8 Distribution of SDQ Scores: MCS Sweep 4 (Age 7) . . . 125
4.9 Distribution of SDQ Scores: MCS Sweep 5 (Age 11) . . . 125
4.10 Distribution of Parent Behaviours: MCS Sweep 2 (Age 3) . . . 129
4.11 Distribution of Parent Behaviours: MCS Sweep 3 (Age 5) . . . 129
4.12 Distribution of Parent Behaviours: MCS Sweep 4 (Age 7) . . . 130

5.1 NLSCY Survey Design . . . 158
5.2 NLSCY Survey Structure . . . 161
5.3 Distribution of Cognitive Scores: NLSCY Cycles 2-6 . . . 174
5.4 Distribution of Non-Cognitive Scores: NLSCY Cycle 3 (Ages 4/5) . . . 180
5.5 Distribution of Non-Cognitive Scores: NLSCY Cycle 4 (Ages 6/8) . . . 181
5.6 Distribution of Non-Cognitive Scores: NLSCY Cycle 5 (Ages 8/9) . . . 181
5.7 Distribution of Non-Cognitive Scores: NLSCY Cycle 6 (Ages 10/11) . . . 182
5.8 Distribution of Parental Investment Measures: NLSCY Cycle 3 (Ages 4/5) 186
5.9 Distribution of Parental Investment Measures: NLSCY Cycle 4 (Ages 6/7) 187
5.10 Distribution of Parental Investment Measures: NLSCY Cycle 5 (Ages 8/9) 187

B.1 Summary of NLSCY Measures . . . 260

List of tables

3.1 Comparing the Performance of Fit Statistics 85

4.1 Number of Respondents: MCS Sweeps 1–5 104
4.2 Definitions of Included Covariates 109
4.3 Time Fixed Demographic Measures: Unweighted Descriptive Statistics . 111
4.4 Time Varying Demographic Measures: Unweighted Descriptive Statistics 113
4.5 Cognitive Measures Included in the MCS 114
4.6 Measures of Cognitive Ability: Unweighted Descriptive Statistics 117
4.7 Non-Cognitive Measures: Strengths and Difficulties Questionnaire 123
4.8 Measures of Non-Cognitive Ability: Unweighted Descriptive Statistics . . 124
4.9 Parental Investment Measures: MCS Parenting Questions 126
4.10 Parental Investment Measures: Unweighted Descriptive Statistics 127
4.11 Measurement Model: Cognitive Ability - Parameter Estimates 135
4.12 Measures of Non-Cognitive Ability: SDQ Scale Reliability 137
4.13 Measurement Model: Non-Cognitive Ability - Parameter Estimates . . . 138
4.14 Parental Investment Measures: EFA Measures of Fit 140
4.15 Parental Investment Measures: Factor Structure 141
4.16 Measurement Model: Parental Investment - Parameter Estimates 143
4.17 Structural Model: Parameter Estimates 144
4.18 Structural Model: Parameter Estimates Initial Period Covariates 147
4.19 Structural Model: Parameter Estimates Covariates for Investment 148

5.1 Number of Respondents: NLSCY Longitudinal Cohort 162
5.2 Definitions of Included Covariates 167
5.3 Time Fixed Demographic Measures: Weighted Descriptive Statistics . . . 169
5.4 Time Varying Demographic Measures: Weighted Descriptive Statistics . . 170
5.5 Cognitive Measures Included in the NLSCY 171
5.6 Measures of Cognitive Ability: Weighted Descriptive Statistics 173
5.7 Non-Cognitive Measures: NLSCY Behaviour Scales 178
5.8 Measures of Non-Cognitive Ability: Weighted Descriptive Statistics . . . 179
5.9 Parental Investment Measures: NLSCY Parenting Scale Questions 184
5.10 Parental Investment Measures: Weighted Descriptive Statistics 185
5.11 Covariate Parameter Estimates: Cognitive Ability 193
5.12 Measures of Non-Cognitive Ability: Scale Reliability 195
5.13 Measurement Model: Non-Cognitive Ability — Parameter Estimates . . 197
5.14 Parental Investment Measures: EFA Measures of Fit 198

5.15 Parental Investment Measures: Factor Structure 199
5.16 Measurement Model: Parental Investment - Parameter Estimates 200
5.17 Structural Model: Parameter Estimates 202
5.18 Structural Model: Parameter Estimates Initial Period Covariates 205
5.19 Structural Model: Parameter Estimates Covariates for Investment 206

B.1 Cognitive Measures: Covariate Parameter Estimates 284
B.2 Parental Investment Measurement Model: Parameter Estimates 285
B.3 Measurement Model: Non-Cognitive Skills - Parameter Estimates 286
B.4 Structural Model: Parameter Estimates Non-Cognitive Ability 287
B.5 Structural Model: Parameter Estimates Cognitive Ability 288

List of Abbreviations

BAS	British Ability Scale
BAS–II	revised version of BAS, the British Ability Scales–2nd Edition
BBCS	Bracken Basic Concept
BBCS–R	Bracken Basic Concept Scale–Revised
BSRA	Bracken School Readiness Assessment
CAI	computer assisted interviewing
CANTAB	Cambridge Neuropsychological Test Automated Battery
CAT	Canadian Achievement Test
CAT–2	Canadian Achievement Test, Second Edition
CBCL	(Achenbach's) Child Behavior Checklist
CFA	confirmatory factor analysis
CFI	Comparative Fit Index
CLS	Centre for Longitudinal Studies
CPI	consumer price index
DAS	Differential Ability Scales
EAL	English as an Additional Language
ECD	early childhood development
EFA	exploratory factor analysis
EVIP	Échelle de vocabulaire en images de Peabody
FAST	Families and Schools Together
FSP	Foundation Stage Profiles
HRSDC	Statistics Canada and Human Resources and Skills Development Canada

IRT	item-response-theory
LFS	(Canadian) Labour Force Survey
MCS	Millennium Cohort Study
MPlus	a statistical software for structural equation modelling
MSD	Motor and Social Development
NFER	National Foundation for Education Research
NHS	National Health Service
NLSCY	National Longitudinal Survey of Children and Youth
NPHS	National Population Health Survey
OCHS	Ontario Child Health Study
OECD	Organisation for Economic Co-operation and Development
PCA	principal components analysis
PiM	Progress in Maths
PMK	Person Most Knowledgeable
PPVT	Peabody Picture Vocabulary Test
PPVT-R	Peabody Picture Vocabulary Test — Revised
RDC	Research Data Centres
RMSEA	Steiger-Lind root mean squared error of approximation
SDQ	Strengths and Difficulties Questionnaire
SES	socioeconomic status
SEM	structural equation modelling
SRMR	standardised root mean square residual
SSHRC	Social Sciences and Humanities Research Council
STATA	a general purpose statistical software package
TLI	Tucker Lewis Index
UK	United Kingdom
US	United States of America
WLSMV	robust weighted least squares mean variance adjusted estimator
WPPSI–R	Wechsler Preschool and Primary Scale of Intelligence–Revised

CHAPTER 1

Introduction

There is extensive evidence to suggest that measures of cognitive and behavioural development, taken in childhood, are highly predictive of a variety of outcomes in adulthood (Carneiro, Crawford, & Goodman, 2007; Cunha, Heckman, Lochner, & Masterov, 2006; Feinstein, 2003; Schoon, 2010; Todd & Wolpin, 2003, 2007). While early investment in child development is now a consistent feature of education policy, it is increasingly evident that gaps in childhood abilities emerge long before the start of formal education and that family circumstances and parental behaviour play a considerable role in the ongoing development of these skills (Ermisch, 2008; Kelly, Sacker, Del Bono, Francesconi, & Marmot, 2011; Schoon, Jones, Cheng, & Maughan, 2012). Consequently, identifying and promoting the precise parental inputs that facilitate childhood skill development is a key component of research and social policy alike. To better understand how parenting influences the joint evolution of these skills, this thesis defines a theoretical framework which models the effect of parental investment on the developmental trajectories of cognitive and non-cognitive skills in primary school children. Using nationally representative, longitudinal survey data from the UK and Canada, this framework serves to estimate skill formation empirically in two contexts.

Investigating the relationship between parental investments and child development is not a new field of study; research spanning the fields of education, economics, sociology, and psychology has found that parental time, attitudes, income and education all help determine children's outcomes (e.g. Guo & Harris, 2000; Leibowitz, 1974; McLoyd, 1998; Paulson, 1994). Though each field has a distinct methodological approach, findings from all disciplines suggest that, in addition to influencing cognitive ability, parental inputs also shape children's behaviour, social skills and emotional development. These outcomes all fall under the umbrella of what economists refer to as *non-cognitive skills*[1] and there is substantial evidence that these skills play their own unique role in the development of

[1] Chapter 2 of this thesis presents a comprehensive definition of *non-cognitive skills*. For now, they can be thought of as including: social skills, emotional skills and other behavioural measures.

cognitive skills (for details, see Cunha et al., 2006). While the importance of non-cognitive skills is well-established, models of child development often focus on cognitive outcomes; meanwhile, non-cognitive outcomes are somewhat, but not completely, neglected. In this thesis, I present a model which includes both cognitive and non-cognitive outcomes.

A variety of inputs must be considered when modelling the development of cognitive and non-cognitive ability. Much of the research into the determinants of childhood skill development focuses on socio-economic factors such as family income, parental education and parental employment status. This existing research aims to identify parenting factors that either explain or mediate the effects of poverty (Berger, Paxson, & Waldfogel, 2009; Brooks-Gunn & Markman, 2005; Ermisch, 2008; Kelly et al., 2011; McCulloch & Joshi, 2002). Often this research finds that poverty predicts certain parenting behaviours — either the presence of negative behaviours or the absence of positive ones — which, in turn, correlate with poor outcomes in children. Even when research is not specifically studying the impact of socioeconomic status (SES), it tends to include family income within the larger set of measures designed to capture parental input (e.g. Cunha, Heckman, & Schennach, 2010; Todd & Wolpin, 2007). By contrast, very little research has been devoted to distinguishing between the effects of SES and those of parenting behaviours, and even less to identifying distinct types of parenting factors. Separately measuring the effect of financial inputs from those of parenting behaviours is critical for identifying targeted policies which apply across the socio-economic spectrum. Similarly, when designing public policy, it is important to distinguish between different types of parenting behaviours. For example, behaviours promoting social skills are not necessarily the same as those promoting numeracy, but these behaviours are often combined — alongside measures of SES — under the umbrella term 'parental investment'.

To address these gaps in the literature, my thesis builds on an existing methodological framework of childhood skill development to allow for the identification of multiple types of parental investment and to examine the unique impact from each of these types of investment. Specifically, I draw on literature from psychology and education to identify multiple parenting factors, beyond financial resources, and include each of these factors as separate inputs in a modified version of the economic skill formation model originally presented by Cunha and Heckman (2007, 2008).

Since this model is designed to distinguish between types of parental input, it allows for more precise estimates of the effect that various parental behaviours have on child development, as well as the identification of the specific periods of childhood during which these investments might be most effective. Such findings offer insights into both the content and the timing of interventions to help at-risk children.

To establish the capacity of my proposed model to measure the role that parental inputs play in childhood skill formation, this thesis applies the model to data from two contexts: the UK and Canada.[2] The results from these two countries yield insights into the nature of skill formation in each context as well as demonstrating how the model can be adapted to existing data. When presenting each empirical application, I discuss the characteristics of the available data, and provide further details on how the findings determined from this data fit within the context of my model and the existing literature.

1.1 Research Aims

This thesis aims to bridge the literature from the fields of economics, psychology and education in order to advance our understanding how parental behaviours shape children's development trajectories. Each of these three fields offers unique insights into how parental inputs influence childhood skill formation. In addition to consolidating the theoretical and empirical findings from the three disciplines, this thesis pays special attention to differentiating between various parental inputs to child development and separating the impact of household resources from the direct effect of parenting behaviours. As such, this thesis seeks to answer three key research questions:

1. **How can the literature from psychology and economics be consolidated within the field of education to form a theoretical framework to explain parents influence the development of cognitive and non-cognitive skills?**
2. **Can existing models of skill development be improved in order to distinguish between parental behaviours and socio-economic resources?**
3. **What can empirical estimates from a variety of contexts tell us about the role that specific parenting constructs play in childhood skill formation?**

In response to the three questions raised above, this PhD thesis has three key objectives. First, I propose a theoretical framework to explain how parents influence childhood skill development. This framework uses existing empirical and theoretical research to propose a model which separates parental behaviours from family resources, thereby developing a more comprehensive understanding of how both these factors lead to the development of later skills. My second objective is to show how this framework can be applied to existing data to provide relevant estimates of the effect that parental inputs have on the development of specific types of skills. Estimating the model using data from two countries allows me to explore the ways in which development varies between different educational and policy contexts. Finally, I aim to discuss the implications of these empirical findings for public policy and for future research.

[2]While this thesis is not a comparative study,these two countries allow me to examine how skill development might differ in two countries which have similar levels of economic development and strong welfare states but differ in terms of social mobility and socio-economic gaps in attainment. More information about the motivations for choosing these countries will be presented later in this thesis.

1.2 Overview of the Thesis

As outlined above, this thesis adapts the skill formation model, originally introduced by Cunha and Heckman (2007, 2008), to create a methodological framework which measures the role that multiple types of parental investment play in the development of cognitive and non-cognitive skills. This modelling approach allows me to distinguish between family characteristics and specific parenting behaviours and to isolate the effect that each of these inputs has on the development of specific skills. To demonstrate the suitability of my proposed model, I apply the empirical framework to data from two large-scale longitudinal surveys. In addition to showcasing how my model can be adapted to multiple contexts, these empirical studies also provide quantitative and policy-relevant evidence on how various parental inputs facilitate skill development.

Organisation and Structure

This thesis is divided into six chapters. Following the introduction presented in this chapter, Chapter 2 presents an extensive review of the relevant literature which explores the theoretical approaches and empirical findings from the fields of economics, psychology and education. I begin this chapter by clarifying some key terms to establish a common language to unite the literature from these three fields. This chapter continues with a review of the existing empirical findings and a detailed discussion of the theoretical frameworks employed by economists and psychologists to model skill formation and parenting. Each of these sections examines the implications that research from the given discipline has for modelling the relationship between parenting and childhood skill development. The chapter concludes by explaining how the field of education provides the methodological space to unify these theoretical approaches and the practical context within which it is possible to collect data and pursue further research on skill formation.

Building on this literature, Chapter 3 introduces and defines an updated model of skill formation that can be used to measure the role that parental investment plays in the trajectory of childhood skill development. Specifically, I contribute to the existing research by using both data-driven and theoretical approaches to identify multiple factors of parental input. I then modify the methodological framework of Cunha and Heckman (2007, 2008) in order to examine how each of these types of investment impacts skill development — critically, parental behaviour and family socio-economic resources are included as separate inputs in the model. Included in Chapter 3 is a detailed discussion of how this theoretical model can be used to estimate skill formation using existing survey data. This contains details regarding the data requirements and methodological considerations of empirically estimating the proposed model, as well as an identification strategy for defining parental input factors within the data.

To demonstrate the validity of my empirical approach, and to produce UK-specific estimates, Chapter 4 applies the empirical framework to a sample (n = 8,379) from the UK Millennium Cohort Study (MCS), a longitudinal cohort survey that follows a nationally representative sample children born in the UK in 2000–2001. Applying my model to the MCS allows me to compare my findings to the recent analysis of Hernández-Alava and Popli (2017). Their study applies a modified version of the skill formation model to MCS data, but only includes observations from birth to age 7. Using my modified version of the skill formation model, this thesis extends the analysis to age 11. While the overlap between the four periods studied by Hernández-Alava and Popli (2017) and the first four periods included in my model allows me to compare my findings with existing estimates, extending my analysis a further period, through to age 11, allows me to measure the developmental trajectories through the end of primary school (and the start of puberty.)

The empirical application presented in Chapter 4 uses measures of specific parent-child activities to identify three separate latent parental inputs.[3] I find that each parental input has a unique impact on the development of both cognitive and non-cognitive skills. Using these three parental inputs, along with measures of socio-economic status, the empirical study is able to provide estimates which capture the effect that parenting behaviours have on skill development, isolated from the effect of family characteristics. To accomplish the joint goals of establishing model validity and providing relevant estimates, the chapter includes an introduction to the MCS data, discusses the estimation strategies used and presents the model estimates.

As the final substantive chapter of this thesis, Chapter 5 applies the proposed framework to Canadian data to demonstrate how this modelling technique can be applied to another context. This second application of the model uses data from a sub-sample (n = 1,234) of the National Longitudinal Survey of Children and Youth (NLSCY) to measure the relationship between parenting and skill development within a Canadian context.[4] Not only does applying the model to Canadian data lend further support for the validity of my proposed methodology model, but it also contributes to the existing literature by estimating skill formation in a country that is known to have less income inequality than either the UK or the US.

Along with estimating the proposed methodology in a different national context, the analysis presented in Chapter 5 also demonstrates the ability of the model to examine another set of parenting inputs. Unlike the UK study, which uses the model to distinguish between various parent-child activities, this Canadian analysis uses parental investment measures that focus on underlying theoretical parenting constructs — such as parental

[3]The three latent parental inputs identified in the MCS data are *literacy activities*; *parent child interactions*; and *academic activities*.

[4]The NLSCY is a longitudinal cohort study which identified a nationally representative sample of Canadian children in 1994 and follows them to adulthood.

warmth and discipline style. Using this approach, I identify three separate latent parental inputs in the NLSCY data.[5] I find that, for every period in the model, each type of parental input has a unique influence on development. In line with Chapter 4, the Canadian analysis presented in Chapter 5 begins with a detailed examination of the relevant data, methodological considerations and estimation strategy; the chapter then presents the estimates obtained from the model and explains how these can add to the findings from the UK analysis.

The thesis concludes with Chapter 6, which not only summarises the findings from the two empirical applications presented in Chapter 4 and Chapter 5, but also situates these empirical findings within the literature presented in Chapter 2 and outlines how they confirm the theoretical framework presented in Chapter 3. This overarching review allows me to discuss my scholarly contributions (both empirical and methodological) alongside the policy implications that stem from my proposed model and its empirical findings. Included in this discussion is a review of possible future work which builds on my research alongside a careful consideration of the limitations of the present study.

[5]The three latent parental inputs which are identified are *positive interaction*; *ineffective parenting*; and *consistent parenting*.

CHAPTER 2

Literature Review

This chapter reviews the growing body of research on the role that parents play in the development of their children's skills and abilities. As this research spans multiple academic disciplines, I not only review the empirical findings, but also outline the pertinent theoretical frameworks. For readers familiar with a given field, some of the information in this chapter might appear overly simplistic and contain details considered to be general knowledge in the field. Though such rudimentary concepts would usually be omitted from a literature review, this thesis proposes a methodological framework that draws on the theoretical perspectives from several fields, and it is important to have a solid grounding in all of the underlying concepts.

The chapter is organised in the following manner. To begin, Section 2.1 clarifies the relevant terminology to ensure consistency for the rest of this thesis. Next, Section 2.2 reviews the existing methodological approaches for measuring skill development and draws on all disciplines to summarise the empirical findings. Then, Section 2.3, and Section 2.4 review the theoretical frameworks for skill development from the fields of economics and psychology respectively. Section 2.5 discusses how the theoretical frameworks from psychology and economics are used in the context of education research. To conclude, Section 2.6 outlines the implications of the literature discussed on this thesis.

2.1 Clarifying Relevant Terminology

This section outlines the terminology I use to describe childhood *skills and abilities*, and *parental inputs and investments*. As the exact meaning of these terms varies between disciplines, it is important to clarify how they are used in the context of this thesis. Special attention is paid to areas with conflicting definitions between academic fields.

Defining Skills and Abilities

The present study follows the convention used in labour economics where the terms *skill* and *ability* are considered synonymous.[6] The interchangeability of these terms in economics is a fairly recent development, with older literature assuming abilities were innate, while skills were acquired. In developmental psychology, a distinction between the terms is sometimes used with *skills* describing acquired proficiency on specific tasks, while *abilities* are more general, enduring traits, that are thought of as relatively fixed.[7] This thesis uses both *skill* and *ability* to describe traits that evolve over time and are the product of a child's environment, genetics, parental investment and formal education.

There is considerable evidence that numerous skills are required for success in life and that these skills develop over the course of childhood (Carneiro & Heckman, 2003; Cunha et al., 2006; Phillips & Shonkoff, 2000). Building on economic theories of human capital, the theoretical framework that forms the basis of my research divides skills into the two categories of *cognitive* and *non-cognitive* skills.[8] Though this differentiation of skills stems from economics, these two types of skills also align with constructs widely used in the fields of psychology and education. Below I define these two types of skill, and how they relate to my research.

[6] Using these terms in this way differs from the way they are usually defined by educationalists. As my proposed model relies heavily on theoretical perspectives from economics, using this particular definition allows for comparability between my findings and existing use of similar models.

[7] This differentiation between skills and abilities was introduced by Fleishman (1967) with a more generalised examination of the concepts provided by Fleishman and Bartlett (1969).

[8] See Cunha et al. (2006) for an introduction to this literature and Heckman and Mosso (2014) for a more recent review of research this area of study.

Cognitive Ability

Within this thesis, cognitive ability refers to an individual's mental capacity in areas such as perception, memory, problem solving, abstract thinking and reasoning. Measures of cognitive ability are perhaps the most researched predictors of academic and labour market success, with research consistently finding high predictive power of cognitive ability for a multitude of outcomes (e.g. Davies, Janus, Duku, & Gaskin, 2015; Duncan et al., 2007; Duncan & Murnane, 2011; Hornung, Schiltz, Brunner, & Martin, 2014; Murnane, Willett, & Levy, 1995; Sabol & Pianta, 2012). For the purposes of measuring the role that parental investment plays on child development, cognitive ability is assumed to be an aggregate measure that captures ability across multiple broad cognitive processes. This conceptualization of cognitive ability is deeply rooted in psychology and the applicable psychological theories are examined in Section 2.4.

Non-Cognitive Ability

In the theoretical framework and empirical analysis that follow, I use the term *non-cognitive ability* to describe the set of abilities that are sometimes referred to as "soft skills." More specifically, non-cognitive abilities include attributes such as personality, emotional intelligence and social skills. There is increasing evidence that these non-cognitive traits are not only predictive in their own right, but also impact the effectiveness of early cognitive ability in translating to positive future outcomes. The research on various non-cognitive skills is extensive and shows that multiple traits are predictive of success. For example, recent research finds strong predictive power associated with attention skills (Brennan, Shaw, Dishion, & Wilson, 2012; McClelland et al., 2007; Spira & Fischel, 2005; Washbrook, Propper, & Sayal, 2013), social-skills (Agostin & Bain, 1997; Arnold, Kupersmidt, Voegler-Lee, & Marshall, 2012; Hartas, 2011; Ladd & Price, 1987; McClelland, Morrison, & Holmes, 2000; Welsh, Parke, Widaman, & O'Neil, 2001), and personality traits (Almlund, Duckworth, Heckman, & Kautz, 2011; Laidra, Pullmann, & Allik, 2007).[9] In the context of examining the role of parenting on child development, non-cognitive ability is assumed to be a general measure that captures a multitude of these individual traits. Precisely identifying and measuring this set of traits requires an understanding of the psychology literature on personality, and emotional intelligence. The relevant aspects of this literature are reviewed later in this chapter, in Section 2.4.

[9] See Borghans, Duckworth, Heckman, and ter Weel (2008) and Almlund et al. (2011) for comprehensive surveys of the literature on non-cognitive abilities.

Defining Parental Input and Investment:

In the theoretical framework presented in Chapter 3, *parental input* and *parental investment* are used interchangeably to describe a broad set of interactions between parents and their children. Unlike cognitive and non-cognitive skills, where the use of relatively precise definitions is a necessary condition for identifying a unifying theory of skill development, parental investment can be defined less restrictively within this framework. This expansive definition of parental input allows the empirical work to capture the many types of parental behaviours that influence a child's development. To ensure that the theoretical framework accounts for this flexible definition of parental investment, special considerations are taken when modelling parental input. A detailed explanation of these considerations is presented in Section 3.3.

The main motivation for defining parental investment to include anything ranging from allocation of parental time, to parenting style, to the specific parenting practices is the increasing evidence that a wide variety of characteristics and behaviours determine how parents shape their child's development. Across the fields of economics, developmental psychology and education, there is substantial research into the ramifications that these various aspects of parenting and home life have on a child's development. In developmental psychology, this research discusses parenting style, theories of development and different parenting behaviours that promote positive outcomes (examples include: Baumrind, 1966, 1978; Bergeman & Plomin, 1988; Bradley, Caldwell, & Rock, 1988). Similarly, economists have long focused on the types of investments that parents make in their children's development of cognitive abilities (notable examples include Becker & Tomes, 1986; Heckman, 2006; Leibowitz, 1974; Todd & Wolpin, 2003). In education, literature focuses on parenting factors that determine school readiness, with particular attention paid on children from low-income families (see for example Davies et al., 2015; Hartas, 2011; Paro & Pianta, 2000). Each of these bodies of work provides valuable insight into the types of parental inputs to include in studies which estimate skill development and how to design suitable modelling strategy that is able to capture these various types of inputs. Section 2.3, Section 2.4 and Section 2.5 provide the specific details on the way parental inputs are defined within each field and how the findings from this literature extend to the model of skill development and to the empirical applications of this model that are presented in this thesis.

2.2 Existing Empirical Research

Though each field uses its own theoretical framework to describe the mechanisms through which parents influence their children's development, similar empirical questions are asked across these different strands of research. This section introduces some key empirical findings to summarise the empirical literature without focusing on which specific discipline the findings emerge from. This is especially important as it allows me to compare the findings from various disciplines in one place, without having to revisit other disciplines each time a new theoretical model is introduced. The findings presented below are only a very small subset of the larger empirical literature on skill formation. I have chosen to only present this small subset in this section in order to identify key themes, but other empirical findings are referenced throughout this thesis.

The Predictive Power of Early Skills (Why do Skills Matter?)

The nature of skills and abilities in the early years is a well-studied topic. The driving force of this research is the assumption that early skills will provide a stepping-stone for a successful path through formal education, which will, in turn; result in positive outcomes in adulthood. Research on the factors which predict this type of success can be found in a variety of fields and this research has found a myriad of predictors of future success (e.g. Bull, Espy, & Wiebe, 2008; Carneiro et al., 2007; Cerda, Im, & Hughes, 2014; Duncan et al., 2007; McWayne, Fantuzzo, & McDermott, 2004; Pagani, Fitzpatrick, Archambault, & Janosz, 2010).

Although the general interest in childhood skills is driven by the evidence linking these skills to adult success, much of the research on skill development focuses on school-readiness. This research is motivated by the joint desires of promoting effective practices in schools and justifying the large public spending on education, which makes up a large part of government budgets. Such research examines the ability of cognitive measures taken in early childhood to predict academic success in the first few years of school and in turn on later development. Duncan et al. (2007) apply regression models to six longitudinal data sets in order to assess whether cognitive ability measured at school entry is predictive of academic achievement at ages 10–11. The authors find that preschool cognitive ability has substantial predictive power for early school achievement, and in turn this predicts academic success at the end of elementary school. Similarly, Davies et al. (2015) use a sample of 45,000 Canadian school children to find that cognitive ability measured prior to the start of first grade is strongly predictive of test scores on standardized mathematics and reading exams administered in the third grade.

The persistence of ability is not limited to cognitive ability with Feinstein (2000) finding that along with cognitive measures, behavioural attributes in early childhood are able to predict later qualifications, educational attainment and earnings. Other research in this field finds strong predictive power associated with attention skills (McClelland et al., 2007; Spira & Fischel, 2005; Washbrook et al., 2013), social-skills (Agostin & Bain, 1997; Arnold et al., 2012; Hartas, 2012; Ladd & Price, 1987; McClelland et al., 2000; Welsh et al., 2001), and personality traits (Almlund et al., 2011; Laidra et al., 2007). An increasing body of research has shown that when measured along with early childhood cognitive ability, these non-cognitive abilities can predict a larger proportion of the variation in adult outcomes (examples include: Attanasio, 2015; Carneiro et al., 2007). In the educational context, these types of skills can be used to identify children at risk of falling behind. For example, Cerda et al. (2014) contend contend that measuring these skills, along with cognitive ability in kindergarten, provides increased predictive power in identifying children at risk of poor academic performance.

Empirical Studies of Skill Development

It is clear from the previous subsection that early childhood skills matter, but what influences the development of these skills over time? A child's development is influenced by a multitude of factors and it would be impossible to address these all in this literature review. The research strategies used to explore the impact of of each of these inputs on child development can be loosely categorized into either correlational studies or experimental research.

Correlational studies on development have explored the role of a multitude of factors such as family income, maternal employment, parental education,and parenting styles. An extensive review of the determinants of childhood abilities was conducted by Haveman and Wolfe (1995), with Cunha et al. (2006) and Heckman and Mosso (2014) providing more recent findings. Instead of replicating the work presented by these surveys, this section provides a general overview the relevant findings and focuses on what these findings might mean for empirical work. This overview is by no means an extensive review of all the literature on parental investment and skill formation, and the reader is directed to the existing literature reviews for more information.

Though correlations studies are able to find associations between cognitive measures, non-cognitive measures and a variety of inputs, the nature of such studies makes it impossible to determine how changing one input would improve later measures of child ability. The gold standard for assessing the ability to change achievement trajectories is to conduct an experimental intervention study. In such an intervention study, researchers select a group of children and apply a specific intervention that changes some input to development (e.g. parental inputs, family resources or some aspect of the child's education) and then measure how the trajectories of these children differs from the trajectories of children who did not receive such an intervention.

In this subsection, I outline the empirical findings on the role that parenting, SES and formal education play in skill development.

The Role of Parenting

Correlational research into the role of parenting on child development is conducted by obtaining measures of parental inputs and developmental outcomes and assessing the statistical relationship between these measures. Unlike the experimental research described later in this section, correlational studies do not involve any attempt to control or modify the behaviour of parents or children and simply measure existing relationships. This allows for researchers to study a wide variety of parenting behaviours and to see if the observed relationships match the correlations that would be expected under proposed theoretical frameworks. While this makes it possible study how parents and children interact without altering their behaviour, it is poorly suited in determining if specific parental inputs are causal in development or simply correlated with a given skill.

The literature contains an expansive set of parental inputs, and a full review would be beyond the scope of this chapter. The reader is again directed to the literature reviews outlined above. There are however, several specific types of parenting inputs are particularly relevant to this thesis. These measures capture three aspects of parenting: the time that parent spends with their children; the types of parenting behaviours that a child is exposed to; and measures of the child's home environment.

The first of these parenting inputs, is the amount of time that parents spend with their children, this can either he measured as a general measure of the time spent together, or as a measure of the time spent engaging in particular behaviours. Del Boca, Monfardini, and Nicoletti (2017) find that as parental time increases there are corresponding increases in both cognitive and non-cognitive outcomes. In a similar study, Yeung, Linver, and Brooks–Gunn (2002) examine parental time inputs alongside measures of the family's home environment and also find that parental time is able to predictive both cognitive and non-cognitive outcomes in 3 and 5 year olds.

The second relevant subset of correlational literature is the research which studies the relationship between parenting behaviours and various childhood skills. This literature can either examine specific parenting behaviours or alternatively can study measures of parenting style which are used to describe the general patterns of parent child interaction. The evidence on parenting style is mixed, in one study examining parenting style, Dooley and Stewart (2007) found that there is no consistent effect of parenting style on youth outcomes, but in another study, using data from the UK, Koo and Chan (2010) find that parenting styles are highly predictive of youth outcomes. The findings regarding parenting behaviours are more consistent: studies by Attanasio, Cattan, Fitzsimons, Meghir, and Rubio-Codina (2015); Bono, Francesconi, Kelly, and Sacker (2016); Cunha and Heckman (2008); Hsin and Felfe (2014); Kelly et al. (2011) and Todd and Wolpin (2007) have all found a positive relationship between parent-child interactions and both cognitive and non-cognitive outcomes. The magnitude of this effect varies depending on the behaviour studied.

Similar findings have resulted from experimental studies on the role that parents play in development, psychologists have also found that parental inputs play a critical role in early skill development. For example, Love et al. (2005) conduct a large-scale review of the American Head-Start Program and found that general measures of parent supportiveness, along with measures of specific parent-child interaction are both predictive of that differences in primary school outcomes. This study included measures of parent-child interactions, assessment of socio-emotional skills (CBCL) in children and cognitive development (PPVT).

The Role of SES

It is commonly accepted that children from poorer backgrounds score lower on measures of cognitive ability. In Section 2.3 I will explore some of the mechanisms through which low-SES impinges on development. The empirical findings relating to the impact of SES on child development are broad, and I direct the reader to Cunha et al. (2006) and the Supplemental Appendix provided by Heckman and Mosso (2014) for an extensive review of these findings. As defined in Section 2.1, this thesis focuses on specific parenting behaviours and does not intend to directly measure the influence that SES has on skill development. While it will be necessary to control for the effect of SES, the details will be presented alongside the empirical analyses in Chapter 4 and Chapter 5. For now, it suffices to note family resources and other measures of SES play an incredibly large role in children's development.

Of specific interest to this thesis is the literature which aims to determine if parenting behaviours are able to mediate the well-known effects that poverty or other deprivation have on skill development. For example, Kiernan and Mensah (2011) show that *positive parenting*[10] corresponds with higher levels of school achievement, across socioeconomic classes and that for low-SES children, this type of behaviour is especially useful and serves to partially mediate the effects of disadvantage. A similar study by Kiernan and Huerta (2008) found that in 3-year olds, the difference in parenting practices between rich and poor households was able to explain more than half of the relationship between economic hardship and childhood cognitive outcomes. An analysis by Mistry, Biesanz, Chien, Howes, and Benner (2008) that parenting practices have a similar mediating effect on the negative relationship between household measures of SES and childhood measures

[10]The authors created a parenting index based on measures which they identified as positive parenting behaviours (e.g. takes child to the library) and positive disciplinary styles (e.g. not yelling at the child).

of aggressive behaviour and cognitive ability.

Alternatively, some researchers chose to use SES as part of a larger combined measure of parental input. Studies by Cunha and Heckman (2008), Todd and Wolpin (2007) and Cunha et al. (2010) have all included measures of family income as part of a larger parenting construct, which they then link to various childhood outcomes. Although this approach is able to show that SES is linked to child development, it is unable to differentiate between the impact of SES and that of the parenting behaviours themselves. This limitation is one of the motivating factors behind the model I present in this thesis, which aims to separately measure the impact of SES and that of other parental inputs.

From all of these studies, there is increasing evidence that some of the SES-based gaps in cognitive and non-cognitive outcomes can be partially explained by the differing parenting practices from across these groups. The mechanisms behind this are varied, and are reviewed in detail in Section 2.2.

The Role of Schools

The final input to skill development that I address in this literature review is formal education. Although the model that I propose is designed to examine parental inputs, it is important to note that much of the existing literature on skill formation is taken from studies which examine various aspects of the formal education system. Research has shown that measures of school quality are able to predict differences in cognitive ability, child behaviour and adolescent measures of psycho-social adjustment (Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993; Duncan & Magnuson, 2013; Duncombe, Havighurst, Holland, & Frankling, 2012; Poulou, 2015, 2017). I have chosen not to present more of the specific empirical findings in this chapter as the modelling approach I define below does not include measures of formal education. On page 29, I discuss why I am able to omit measures of formal education from my model.

2.3 Theoretical Concepts: Economics

Within the field of economics there is substantial research that attempts to identify the childhood predictors of adult success and the role that various parenting behaviours have in determining academic and labour market outcomes. This literature falls into two broad categories. The first type of research, which has already been examined in Section 2.2, identifies relationships between various early childhood inputs and later developmental outcomes observed within the data. The second type of research falls under the umbrella of human capital theory. In this strand of research, economists use consumer choice theory to model how parents allocate their resources between their own consumption and investing in the development of their children. Although economists will be well grounded in the neoclassical theories on which human capital models are based, an overview of the relevant concepts is presented for researchers from other disciplines.[11]

Using Human Capital Theory to Model Development

In economics, *capital* describes the assets used in the production of goods or services. For example, both bicycles and delivery vans are types of capital used by courier companies. Similarly, *human capital* refers to the "set of skills [possessed by workers] that can be 'rented out' to employers" (p. 331 Ehrenberg, 2017). Using the same example, the ability to ride a bike or drive a delivery van are both types of human capital that a worker can 'rent' to an employer. To provide delivery services, a company needs both physical capital (i.e. bike or delivery van) and human capital (i.e. delivery person).

Human capital theory proposes that just as firms choose the amount of physical capital that maximises profits, individuals choose the amount of human capital that optimises their well-being (utility). Utility is derived from the enjoyment of leisure time and the consumption of goods and services that are purchased using the wages paid by employers to 'rent' the use of an individual's time. If a worker possesses more skills, employers are willing to pay higher wages to 'rent' their time.[12] Using the example above, the courier company would choose between purchasing a van or a bicycle for deliveries and the individual would choose between learning to ride a bike or to drive a van.

The optimal level of investment for the firm depends on the price of different technologies, the cost of inputs such as labour and physical resources and the market price of the good they are producing. Similarly, the optimal level of investment in human capital depends on the individual's capacity to learn new skills, the value they place on leisure and the costs associated with any investment in skill development. While firms

[11] A full review of neoclassical economics is beyond the purview of this thesis. See Mankiw (2015) for an introduction to neoclassical theories of utility maximisation, rational preferences and their role in the labour market. Mathematical details of these models are presented by Mas-Colell (1995).

[12] Traditional supply and demand models can be used to explain the determinations of these wages. Readers unfamiliar with these economic concepted are directed to Mankiw (2015) and Dixit (2014).

can increase their physical capital by investing in machinery that allows for more efficient production, it is assumed that individuals can increase their human capital by investing in the development of skills that allow for more efficient production of goods or services. Both these types of investment come at a cost. Firms choose the level of investment that maximises profits given the market price of the good they produce, the cost of inputs and the cost of investment, and people choose a pattern of investment in human capital that maximizes their lifetime utility given the market wages for different skills, the cost of investment in human capital and their own preferences for consumption and leisure.

The parallel between these two decisions extends to the mathematical models that are used to determine the optimal level of each type of investment. Both the profit maximisation decision of the firm and utility maximizing allocation of resources between consumption and skill development for the individual can be explained using a utility maximisation framework. In this type of framework, the 'decision-maker' chooses the set of inputs that provides the best possible outcome, given specific constraints such as budget and prices, and subject to the decision maker's preferences. Though the specific inputs and constraints vary, the mathematical framework for utility maximisation is consistent across applications and used by economists to explain a variety of decisions.

Modelling human capital using utility maximisation frameworks can be traced to the works of Mincer (1958) and Becker (1962, 1964, 1965). Their seminal works show how different levels of investment in human capital can explain the observed distribution of earnings and how household decisions about consumption and leisure can explain the amount of time individuals invest in developing human capital. Ben-Porath (1967) extends these early works to mathematically model how the constraints that an individual faces determine the optimal level of investment in the creation of his or her own human capital.[13] Though Mincer (1958), Becker (1962, 1964, 1965) and Ben-Porath (1967) originally describe adults investing in the development of their own skills, the theory can be extended to cover the decisions parents make regarding investments in their children's development. It is assumed that children have limited control over their environment and that human capital acquisition in childhood is strongly influenced by parental inputs.

Becker and Tomes (1986) formalise the nature of parental investment in children by introducing a model whereby the level of parental input is a function of "utility-maximizing parents who are concerned for the welfare of their children" (p. S1). Parents must allocate their resources between their own consumption and investment in their children's human capital. It is assumed that parents gain utility from the increase in their children's future income that stems from higher human capital. Thus, parents are willing to sacrifice some of their current consumption to invest in their children's development. This trade-off can be captured using the utility-maximisation models described above.

[13]The Ben-Porath model forms the basis of most modern economic models of skill formation.

Applying Economic Production Functions to Skill Formation

A key component of the utility-maximisation framework is the *production function*, which is a mathematical representation of the maximum amount of a good that can be produced by a given level of inputs. Since the ratio between inputs and outputs depends on the type of production technology that a firm has, a production function can be referred to as a *technological process* or as the *technology of production*.

Technological processes were conceived to describe the production of goods and services, but they can also be applied to the production of human capital. Just as a production function is able to explain the ratio of raw materials and labour needed to form a physical good, it can also model how genetic ability, parenting, education and on-the-job training combine to form human capital. For both physical goods and human capital the same basic mathematical framework is used to define the production function.

Using this framework, the production function for any output, Q, is given by:

$$Q = f\left(X_1, X_2, ...X_n\right) \tag{2.1}$$

where $X_1, X_2, ...X_n$ are the inputs to production, also called *factors of production*, and the mathematical function f represents the way in which these inputs combine to create the output Q. The function, f, can take a variety of forms with the exact specification depending on the item being produced, and the level of technology available.

There are several key terms that provide consistent language for discussing the nature of production functions. These are:

- *Substitute:* Two inputs are considered *substitutes* in production if one can replace the other in the production of a good.
- *Complement:* Two inputs are considered *complements* in production if the presence of one input makes the other input more efficient at producing a good.
- *Marginal Product:* The increased output resulting from one additional unit of a given input is known as the *marginal product*.

Early Empirical Studies using Human Capital Production Functions

Using this conceptual framework, Ben-Porath (1967) modelled human capital acquisition with a recursive production function in which the current amount of human capital is the product of all prior investments, along with an individual's innate ability. This innate ability is considered to be a previous endowment that is fixed at birth and varies from individual to individual. The Ben-Porath model is generally accepted as the conventional form for the technology of skill formation and the majority of economic research on skill formation uses variations of this production function.

The first researcher to model parental investment using a Ben-Porath technology was Leibowitz (1974). Using a modified production function, she finds that maternal time spent on instructional activities correlates with increased human capital, as measured by IQ. This research was critical in demonstrating the capacity of technological processes to capture parental investment in childhood human capital formation. Though seminal, this work uses a narrow definition of human capital, as it only examines cognitive ability.

As part of a larger theoretical framework examining the generational persistence of earnings, assets and consumption, Becker and Tomes (1979, 1986) include a modified Ben-Porath production function. They present a multi-generational model, where individuals experience one period of childhood and one period of adulthood. In adulthood, parents allocate their resources among their own consumption, investment in their children's human capital, and transfers of assets to their children. Individuals gain utility not only from their own consumption, but also from the future consumption of their children. This future consumption results from assets received from intergenerational transfer and from the increased earnings resulting from parental investments in a child's human capital.

While the theoretical framework of Becker and Tomes (1979, 1986) is groundbreaking for using both investments in human capital and direct transfers of assets to explain the persistence of wealth across generations, it has several key limitations. The largest limitation is that it defines childhood and adulthood each as a single period. This two-period design prevents the model from capturing specific stages of development and implies that parental investments made at any point in childhood are equally effective. Furthermore, like Leibowitz (1974), the Becker-Tomes model uses a unidimensional measure of human capital. While Becker and Tomes (1986) acknowledge that "human capital takes many forms" they "simplify [the model] by assuming that it is homogeneous" (p. S6). This simplification prevents the model from capturing the true nature of human ability and how different investments might relate to specific skills.

These limitations aside, the work of Becker and Tomes provides a strong theoretical framework for the modern literature on the role of family influence in the development of skills. The full household production model described by Becker and Tomes (1986) captures areas of parental decision making that extend beyond the present thesis and into the factors determining parental behaviour. While these areas are of interest for other reasons, the present thesis aims to analyse the impact of parental behaviours and characteristics on child development and is less concerned with the motivations behind these parental behaviours. More specifically, the present thesis treats these behaviours and characteristics as largely exogenous and is concerned more with the impact of them, not the factors that cause them to exist. For example, the model does not differentiate between a parent who takes their child to the park because they believe the interaction with other children is important for fostering social skills (i.e. parental investment in the child's development) and another parent who does so because they themselves enjoy socialising with other parents at the park (i.e. parental consumption in the form of personal enjoyment of leisure time). Fortunately, within the framework presented by Becker and Tomes, it is possible to focus on one component of their larger model. Thus, while parental motivations influence behaviours, these motivations are not the focus of this thesis and it is possible to examine the role of parental behaviours independently of motivation. This decision to examine behaviours without considering the factors that drive these behaviours is further supported by the psychological theories of parenting presented in Section 2.4.

Contemporary Empirical Studies using Human Capital Production Functions

Building on the early human capital studies described above, many economists have used production functions to model the development of skills in both children and adults. The empirical findings from many of these child-focused studies have already been presented in Section 2.2. Now that I have introduced the concept of production functions, the specific theoretical details of these models can be discussed in the remainder of this chapter. Further details on how features of the model have been specified in various existing studies will be included in the methodological presentation provided in Chapter 3.

As the full range of human capital production functions is beyond the scope of this thesis, I have not revisited these specific studies in this section. However, the relevant details of how models of human capital are defined to measure childhood skill formation are discussed in the following section, along with reference to the pertinent research. For a review of this literature, the reader is directed to the reviews of recent theoretical approaches presented by Heckman and Mosso (2014), Francesconi and Heckman (2016) and Attanasio et al. (2015).

Key Considerations for Contemporary Human Capital Production Functions

Following the seminal work of Becker and Tomes (1986) the economic models used to capture family decision making have evolved substantially. Although these developments extend to all areas of the family choice model, the present discussion focuses on the evolution of Ben-Porath style production functions.[14] While this type of human capital production function is widely used, existing research has differing opinions in regards to three key aspects of the production function. These three considerations are: how human capital is defined within the model, the factors of production that are included in the production function, and the correct specification for the functional form of this production function (Attanasio et al., 2015).

[14] See Aiyagari, Greenwood, and Seshadri (2002) for an excellent explanation of how this type of production function fits within a larger household production model.

Defining Human Capital: Earlier in this chapter, Section 2.1 explores the literature which describes human abilities and skills as a combination of *cognitive skills* and *non-cognitive skills*. The work of Carneiro and Heckman (2003) proposes that both these types of skills are valued by employers and as a result, human capital is a composite of cognitive and non-cognitive abilities. Building on this logic, Cunha et al. (2006) model human capital as a two-dimensional vector of non-cognitive and cognitive skills.

Although many other empirical studies only model the production of cognitive ability, I follow the lead of Cunha and Heckman (2007, 2008) and Cunha et al. (2010), and use a multi-dimensional measure of human capital which includes both cognitive and non-cognitive skills. This modelling approach is supported by the wealth of empirical evidence which shows that a multitude of skills are related to success (e.g. Almlund et al., 2011; Borghans et al., 2008; Cunha et al., 2006; Phillips & Shonkoff, 2000). Further support for this approach comes from theoretical models of child development that are presented in Section 2.4.

In addition to requiring that both types of skills be identified as outputs, it is also important to model skill formation in such a way that both types of skill are estimated simultaneously. This is because there is substantial evidence that these skills interact and build on each other to form skills in the next period.[15] Although the models presented by Cunha and Heckman (2007, 2008), and Cunha et al. (2010) emphasize the importance of estimating these skills simultaneously, this consideration is often neglected in other empirical studies. For example, Attanasio (2015) uses linear models that estimate separate production functions for each type of skill separately. This approach fails to capture the cross-productivity of skills and how this influences developmental trajectories.

Identifying the Factors of Production: The literature discussed in Section 2.2 shows that numerous factors are correlated with childhood skill development. This includes not only financial resources, but also specific parenting behaviours and other family-specific characteristics. Based on this literature, it is clear that there are many parental inputs to skill development, and that each of these different types of inputs has demonstrated the capacity to change a child's developmental trajectory. From Section 2.2 it is clear that many measures of parenting behaviours are simply proxy measures for socio-economic factors, and for this reason it is important to include SES as a separate input in production to avoid this omitted variable bias.

Though there is extensive literature on the inputs to skill formation, the diversity of this research means that there is no prescriptive definition for a comprehensive list of inputs to the production function for human capital. Instead, production functions should be modelled in such a way that they can capture a wide variety of parental inputs with each application of the model defining the inputs based on the available data and

[15] This evidence is discussed later in this Literature Review.

the focus of the study. This approach closely aligns with the existing literature, with the methodological review of Heckman and Mosso (2014) acknowledging the multitude of parental inputs and suggesting a flexible production function. In Section 3.1 I discuss the details of such an empirical strategy, with specific focus on the underlying mathematical framework for estimation.

Specifying the Production Function: Todd and Wolpin (2003) introduce a general framework for modelling a production function for cognitive skills. Within this framework, they interpret a variety of production functions from the literature, with specific focus on the underlying theoretical assumptions of each approach.

Extending on this work, Todd and Wolpin (2007) identify five types of cognitive skill production functions. These are:

- *Contemporaneous specification.* Factors of production are all captured in a single period, and these inputs create the output of cognitive ability in the same period.
- *Cumulative specification.* Inputs to cognitive skill production are from multiple periods. This allows for lagged inputs to contribute to present skill formation.
- *Fixed-effect specifications.* This approach is an extension to both contemporaneous and cumulative specifications, and controls for unobserved endowments.
- *Value-added specification.* In cases where information on lagged inputs is limited, this model is designed to use lagged measures of cognitive ability to partially capture prior inputs. Since lagged measures of skill are used as a proxy for prior investment, this type of model is unable to differentiate between the impact of prior skill and the impact of lagged inputs.
- *Value-added plus lagged inputs specification.*[16] Using both lagged inputs and previous measures of cognitive ability, this modelling strategy is able to separate the impact of prior ability from the impact of prior investment. This captures many of the features of skill formation discussed in the developmental psychology literature, but requires detailed data, collected across multiple periods (Todd & Wolpin, 2007).

To identify the most effective modelling approach, these “alternative model specifications of the production function are compared using a cross-validation criterion” (Todd & Wolpin, 2007, p.127). They find that the *value-added plus lagged inputs* model yields the most reliable results, with estimates from this type of model providing the best forecast of cognitive skill development in cross-validation studies.[17] Although Todd

[16] Todd and Wolpin (2007) also refer to this model as the *augmented value-added model.*

[17] Cross-validation involves estimating the model on subsets of the full sample. Using these estimates, outcomes are predicted for a different subset of the sample, these are then compared to observed values.

and Wolpin (2003, 2007) focus on cognitive skills, the developmental theories underlying their approach apply to all types of human abilities. The assumption that the validity of the *value-added plus lagged inputs* extends to a broader definition of skill is reflected in the literature, with many contemporary models of skill formation using this specification (e.g. Attanasio, 2015; Biroli, 2017; Cunha & Heckman, 2008; Helmers & Patnam, 2011).

This *value-added plus lagged inputs* specification is a recursive model wherein skills are formed over multiple periods and the skills 'outputted' from one period, along with additional factors from the child's environment, become inputs for the next period. If human capital is given by θ_t[18], then this type of production function can be expressed as:

$$\theta_{t+1} = f\left(\theta_t, I_t, X_t\right). \tag{2.2}$$

whereby a child's human capital in the next period, is given by θ_{t+1}, and is a function of their prior skill θ_t; parental investment in the prior period I_t; and observable exogenous measures of family characteristics and socioeconomic status X_t. Using repeated recursive substitutions of skill, this production function can be re-written as a function of an individual's: time invariant family characteristics X, time varying socio-economic status X_{t-1}, ability at birth θ_1, and all past investments $I_1, \ldots I_{t-1}$. This gives the function:

$$\theta_t = m\left(\theta_1, I_1, \ldots I_{t-1}, X, X_{t-1}\right) \text{ for all } t = 1, \ldots, T. \tag{2.3}$$

Readers are directed to the works of Todd and Wolpin (2007) for evidence on the ideal specification for this production function and to the works of Cunha and Heckman (2007), Cunha et al. (2010), and Biroli (2017) for the specifics of empirically estimating this type of model. Relevant findings obtained using this approach have already been discussed in Section 2.2.

[18]At time t, human capital can be represented by the vector $\theta_t = \left(\theta_t^C, \theta_t^N\right)'$ where the stock of latent cognitive skills is given by θ^C and latent non-cognitive skills by θ^N.

Other Theoretical Considerations from Economics

In addition to the technology of skill formation outlined in Section 2.3, the literature from economics includes several other strands of research that have theoretical implications for how to model child development and parenting. These are reviewed below.

Socio-Economic Disadvantage and Skill Development

There is substantial evidence that children from disadvantaged households face multiple factors that impact their skill formation. Section 2.2 has already presented the empirical research linking SES and various developmental outcomes, but it is important to contextualize these findings within the theoretical framework presented above.

Before describing how SES is incorporated within my empirical model, it is important to clarify the scope of my thesis. Specifically, the goal of my study is to understand how skills develop over the course of childhood and how parenting behaviours influence this development. Since SES plays a role in many aspects of children's lives, a full review of the relationship between SES and skill formation would go well beyond what is needed to understand the theoretical models used in this thesis. For a detailed review of how the impact of SES on child development is studied by economists, the reader is directed to Cunha et al. (2006) along with the Supplemental Appendix provided by Heckman and Mosso (2014). More recently, Caucutt, Lochner, and Park (2017) have discussed how these findings can be explained within economic models of human capital formation and parental investment choices. While this expansive literature shows just how far-reaching the impact of poverty is, it also illustrates how difficult it can be to disentangle the direct effect of poverty from the many factors that themselves are associated with poverty but have their own impact on child development.

Thus, while measures of SES are included in the empirical model used in this thesis, I must emphasize that examining the effect of SES on development is not the aim of the present study. Consequently, the discussion of SES below is limited to the theoretical implications that SES has for general models of skill formation. To do this, I examine three separate pathways through which SES can influence a child's development.

Fewer Resources: The first pathway through which SES influences child development is by directly effecting the resources available to children. This results in *omitted-variable bias* whereby the effect of SES is falsely attributed to parenting behaviours that are compromised by lack of resources. The impact on resources can either take the form of reduced levels of the parenting behaviours being examined by the model or the form of reduced levels of other inputs to development. As discussed in Section 2.2, SES has been shown to impact household resources, quality of childcare, availability of nutritious food, school quality, and availability of enriching extracurricular activities. Section 2.2 also discusses how lower levels of these inputs is known to influence developmental outcomes.

The impact of limited resources on development can be explained using the human capital production function. Like any production function, the technology of skill formation assumes a positive marginal product.[19] By design, this positive marginal product will already account for the reduction in parental input correlated with low-SES. Put differently, if the number of books in the home is included as a measure of parental input in the model, and low-SES reduces the number of books a parent can afford to buy, then that reduction will show up in the data as a lower reported number of books. In turn, from the positive marginal product, this lower input will correspond with lower predicted development. What the production function does not capture is the impact of SES on resources not directly measured by the model. For example, if low-SES predicts poor school quality, which is not considered a parental input, then the impact of SES will not be captured by the production function. To address the omitted-variable bias that would arise from failing to account for the impact of SES on development, the production function should be modified to include SES as one of the factors of production.

Changing Parental Behaviour: Beyond limiting a family's resources, research has shown that poverty and other forms of deprivation change how parents interact with their children. If parenting behaviour and SES are correlated, then, even after controlling for the availability of resources, low-SES children may face different inputs from their high-SES peers. More specifically, there is evidence that specific parenting behaviours are correlated with socio-economic status. Consequently, if relevant socio-economic indicators are not separately controlled for, they may be captured as part of the impact of specific parenting behaviours. For example, Hart and Risley (1995)[20] found that low-income children are exposed to a significantly smaller vocabulary than their high income counterparts, and Ermisch (2008) has measured a "strong association between parents' household income and favourable parenting practices"(p. 69). More recently, Kalil (2015) presented research which examined the causes and consequences behind "economically advantaged parents display[ing] more optimal parenting behaviors across a range of domains"(p. 67). From these studies, it is clear that the impact of SES on parenting behaviour is a potential source of omitted variable bias, and detailed controls for socio-economic status along with the parental inputs of interest must be included in my model. If these confounding factors are not controlled for, the model may falsely attribute the direct impact of poverty to the parenting behaviours themselves.[21]

As with scarce resources, the implication that changes in parental behaviour has on the theoretical framework is simply the reduction of specific inputs to the model. As discussed above, by the nature of the production function, the model already captures

[19] A positive marginal product implies that as input rises, so does output.

[20] Though commonly used as an example of the relation between income and parenting, the magnitude of the findings from Hart and Risley (1995) may be smaller than originally reported. Recent work from Sperry, Sperry, and Miller (2018) has failed to replicate the findings reported by Hart and Risley (1995).

[21] See Kalil (2015) and Lareau (2011) for more information on how parenting differs by social class.

any difference resulting from different investment. For example, if low-SES causes parents to spend less time reading to their children, this would be captured by the fact that parents with low-SES report less of this behaviour. Therefore, the link between parental behaviour and SES does not require any further adjustments to the production function.

Changing Effectiveness of Inputs: The final pathway through which SES influences child development is to fundamentally change the effectiveness of a given set of parental inputs. For example, if SES was shown to directly impact a child's innate ability, this would in turn change the way that the child responds to investment over the course of his or her lifetime. This *interaction effect*[22] requires a model which not only controls for SES, but also allows for it to change the way that other inputs promote development.

As discussed in Section 2.2, there are conflicting findings about the exact role that SES plays in changing the nature of cognitive and non-cognitive development. Fortunately, from a theoretical standpoint, it is not necessary to quantify this relationship before specifying the model. Instead, the exact interaction effect can be estimated by the model by including interaction terms for SES and each of the measured inputs. The specific details of this modelling strategy are reviewed in Chapter 3 but for now it is sufficient to note that it is possible to control for this pathway by including measures of SES as predictors for cognitive and non-cognitive ability at all stages of the model.

Modelling SES in this Thesis: Due to the intricate relationship between SES and child development, SES can be included in the technology of skill formation in many ways. Unsurprisingly, there is extensive research aimed at separating the direct effect of poverty from the indirect effect of these poverty related factors. In one such study, Dickerson and Popli (2016) examined data from the Millennium Cohort Study to find that "poverty not only has a direct negative effect on children's cognitive development, but it also has an indirect effect through its adverse effect on parental inputs. (p. 556). This confirms the findings of Dahl and Lochner (2012), Schoon et al. (2012) and Kiernan and Mensah (2011) who all found measurable direct effects of poverty on development that cannot be explained by confounding factors (i.e. parental behaviour, parental educational attainment) alone. These findings are not limited to the UK. Bradbury, Waldfogel, Washbrook, and Corak (2015a) compared the differing home environments of children in the United States, Canada, the UK and Australia to show that in all countries, while parent-child interactions can partially explain income related gaps in skills, there is still a direct effect of SES on development.

Unfortunately, controlling for all mechanisms through which SES influences skill formation is impossible. Even if the above strategies are used to incorporate SES in the production function, SES must be considered when interpreting any empirical findings.

[22] Interaction effects occur when one input makes another input more (or less) effective than it would be on its own. In such cases, the variables' joint effect is larger than the sum of their individual effects.

Separating Parental Motivations from Parental Investment

As discussed earlier in this section, the technology of skill formation is only a small component of a larger theoretical framework for the consumption and investment decisions facing families.[23] Since this thesis focuses on skill formation, this literature review and the empirical estimations that follow do not explicitly examine, or model, the other components of the household utility-maximisation framework. Instead, the production function for skill formation is examined in isolation but understood to exist within this larger framework.[24] In order to examine the production function in isolation, it is assumed that the factors determining the inputs to skill formation do not change the nature of skill formation itself. Put differently, the inputs to production are exogenous.

This assumption of exogeneity is key to defining the role that parental motivations play in the technology of skill formation. Specifically, it implies that the decisions made by parents regarding how to parent their children are not included in the production function for skills. This is because when determining how much skill is produced from a given parental input, it is the amount of an input that matters, not the reason that a parent has for providing the input. While parental motivations and attitudes may directly impact on child behaviour and learning, within the skill formation model this effect can be separated from the effect that specific parenting behaviours have on child development. Put differently, while understanding what drives parental investment is one component of the larger household-maximisation problem, the factors that motivate parents to invest in their children are not considered components of the production function for skills.

To explain why parental motivations are not included in the production function for skill, I briefly revisit the role this function plays in the larger household maximization problem. In this framework, it assumed that parents choose to invest in their children in order to gain utility from their child's outcomes. The skill formation model (production function) is contained within the larger model in order to determine what level of outcome results from a given investment. In turn, the value that parents place on this level of output is determined by parental motivations.

The way in which economists conceptualize these motivations has changed over time. In early human capital production functions, parental investment was assumed to be purely *altruistic*. This altruism was based on the assumption that parents valued their child's future utility (Becker & Tomes, 1986), with this utility being defined by the

[23]Detailed exposition of the household maximization problem is presented by Becker and Tomes (1986), Cunha and Heckman (2007), Cunha et al. (2010), and Attanasio (2015).

[24]It is not uncommon to examine one component of a larger equilibrium model in isolation. For example, for the equilibrium model of production, which determines the level of goods that a company must produce to maximize a firm's profit, the production function which defines how labour and capital combine to form the good is considered to be independent from the supply of these items.

child's own preferences. More recently, instead of being modelled as altruistic, parents have been modelled as *paternalistic*. In paternalistic models, parents value not only their child's utility, but also gain utility from shaping their children's outcomes in a way that aligns with the parent's own preferences (Doepke & Zilibotti, 2017). One example of paternalistic parental investment is the investments that parents make in teaching their children common cultural, religious and social traits (Cunha & Heckman, 2009; Doepke & Zilibotti, 2017). In these cases, parents have specific interest in their children choosing outcomes that align with familial values. While parental paternalism and altruism influence the value that parents assign to a given outcome that results from skill formation, changing the value of the outcome does not change the level of skill that is produced from a given amount of parental input.

Therefore, the preferences driving parental investment are outside the scope of my empirical model, and the production functions, presented in Chapter 3, define parental inputs as exogenous. This exogeneity implies that the model will yield the same results regardless of parental motivations being altruistic or paternalistic.

However, while parental motivations do not influence the impact that specific parental inputs have on a child's development within the model, any policy recommendations stemming from the model should be contextualized within the larger literature which accounts for both altruistic and paternalistic motivations. More specifically, while parents are assumed to care about improved outcomes, they also base their investment decisions on how certain parenting behaviours align with their own set of values. If policy makers recommend parenting practices that violate these values, parents are unlikely to change their behaviour. Thus, policy must be guided not only by empirical findings from the model, but also by the literature regarding parental investment preferences.

Theoretical Implications of Formal Education

Though skill formation is often discussed within the context of household optimization models, children receive inputs to their development from outside the home environment. The largest of these factors is the formal education system, and Section 2.2 has already reviewed some of the research which finds that skills are influenced by school-based factors in addition to family inputs. This research is very critical for identifying areas within the education system that can most effectively promote child development, but for my research I am only concerned with how school-based inputs to development might cause omitted-variable-bias when production functions are used to model the effect that parents have on child development.

In an ideal world, there would be extensive measures of school-quality alongside the measures of parental input required for the model that I use for my empirical studies. If this were the case, it would be possible to define human capital production functions that also include formal education as an input to production. Unfortunately, this type of data is rare, and models of skill formation tend to focus on either school-based inputs or family-based inputs. Models examining various school-level measures as inputs to production have been used to measure the role that formal education plays in child development. These models are known as *education production functions.* For a recent review of education production functions, the reader is directed to Britton and Vignoles (2017) which examines inputs to production including school type, teacher quality and school resources. The models explored by this literature and the empirical results that these models yield, provide valuable insight for policy discussions and help contextualize the findings from skill formation models that are focused on parental inputs.

While it is clear that school-based factors serve as inputs to the production of childhood skills, the growing consensus is that the majority of differences in child outcomes can be explained by factors outside the school system. In one such study, Rasbash, Leckie, Pillinger, and Jenkins (2010) analysed twin data from the UK to find that only 19% of the variation in cognitive outcomes can be attributed to children's experiences in school, with the majority of the remainder being the result of variations at the family level.

As school choice is strongly linked to parental characteristics (i.e. SES) and parenting behaviours, there is likely a strong correlation between the school-level inputs and parent-level inputs. This thesis does not include a review of the school choice literature, as many aspects of this literature extend beyond the realm of my model. I do note that if the quality of educational inputs is driven by a family's available resources, then some of the effect of these inputs can be controlled in human capital production functions by using measures of SES as control variables in the model.

Due to the strong connection between SES, school-choice and school quality, many of the differences in school-level inputs will be partially captured by control variables which capture a child's SES. For those school-level inputs that I am unable to control for, the impact on child development is not large enough to overshadow the returns to parental investment that are captured by the household-level human capital production functions I have described above. That said, it is still important to consider that skill formation does not occur in a vacuum, and therefore any results must be evaluated with a consideration for factors beyond the family that might shape development.

2.4 Theoretical Concepts: Psychology

Although Section 2.2 has already presented some of the correlational and experimental psychology research that identifies various parenting related predictors of cognitive and non-cognitive development, it is important to understand the theories motivating this empirical research. Unlike economics, where the applicable theoretical framework for the present study is mainly drawn from a single strand of research (human capital theory), in psychology the relevant literature comes from multiple strands of research (developmental psychology, psychometrics, social psychology, and cognitive psychology). The important details from the relevant sub-fields of psychology can be loosely divided into three areas. Each of these areas is examined in separate subsections below. To begin, Section 2.4.1 explores how psychological theory can be used to define cognitive and non-cognitive skills. This exploration of theory includes an overview of the psychometric literature that provides a framework for adequately measuring these skills. Next, Section 2.4.2 introduces the theories of child development that inform modern research in developmental psychology. Included in this subsection is a discussion of the implications that each theory has for empirically modelling skill formation. To conclude, Section 2.4.3 examines psychological theories on parenting and discusses how this research can be used to identify the inputs to child development models.

2.4.1 Defining Skills Using Psychological Theories

Earlier in this chapter, Section 2.1 outlined basic definitions of cognitive and non-cognitive skills. Although these definitions allowed for a thorough review of existing empirical work in Section 2.2 and an exploration of the way skill development is modelled by economists in Section 2.3, a more detailed specification of these terms is central to the model developed by the present study.

This subsection examines the psychological literature on defining and measuring cognitive ability, along with research on personality, and emotional intelligence and how inventories of these traits and abilities can be used to measure non-cognitive ability. As presented in Section 2.2, there is substantial literature which links parenting with these two aspects of non-cognitive ability. Similarly, measures relating to emotional intelligence and personality are often included in the school readiness literature from the discipline of education. This thesis therefore uses these two measures as a starting point for specifying non-cognitive ability. As I will explain later in this thesis, the measures used to model non-cognitive ability will likely address more than one psychological construct, and therefore the explanation of the two constructs below should not be considered an exhaustive presentation of all possible non-cognitive skills.

Cognitive Skills

In a great deal of social science research, the term cognitive ability is used interchangeably with the term intelligence. Though this thesis does not specifically consider intelligence as an outcome of child development, measures for various aspects of cognitive ability are often strongly correlated with scores on traditional IQ tests (Schoon, 2010). More importantly, many tests of cognitive ability included in the longitudinal data that is used to study child development are based on intelligence tests originally designed to reflect the way in which psychologists conceptualized human intelligence. While the underlying psychological theories have developed over time, some of the tests still reflect historical aspects of intelligence theory.[25] For this reason, a basic understanding of the psychological theories of human intelligence that forms the foundation for these assessments of cognitive ability is crucial to understanding the types of measures that are included in much of the existing data. Though most psychologists are familiar with modern models of human intelligence and how these models correspond with the cognitive measures used in longitudinal data, an overview of the relevant details is presented below for readers from other disciplines who are unfamiliar with this literature.[26]

The longitudinal data used in this thesis includes cognitive measures that are based upon the *Cattell-Horn-Carroll Theory of Cognitive Abilities* (CHC Theory). This theory consolidates the models of intelligence proposed by Cattell (1941, 1963), Horn and Cattell (1966) and Carroll (1993). These three models are as follows.

- *Cattell Gf-Gc Theory*: Cattell (1941) proposed that there are two distinct types of human intelligence: fluid intelligence (*Gf*) and crystalized intelligence (*Gc*).
- *Horn-Catell Gf-Gc Theory*: Over time, further types of intelligence were added. This resulted in the *Horn-Catell Gf-Gc Model*, which identifies ten distinct abilities (Horn, 1968, 1991; Horn & Cattell, 1966; Horn & Noll, 1997; Horn & Stankov, 1982).
- *Carroll's Three-Stratum Theory:* Building on the idea that multiple cognitive abilities exist, Carroll (1993) presents a hierarchical model where a single '*general*' intelligence factor is formed from a set of eight *broad* cognitive ability factors, which are themselves formed of 70 *narrow* abilities. The *broad* abilities included in this hierarchical model are similar to the types of intelligence in the *Horn-Catell Gf-Gc Model*.

CHC-theory is assumed to reflect features of both Carroll's Three-Stratum Theory and Horn-Cattell's extended Gf-Gc Model. McGrew (1997) presented the first formalized CHC-model, but multiple updated versions have emerged since (Kaufman, Kaufman,

[25] See Snyderman and Rothman (1988) for a review of the history of intelligence testing.

[26] While the full details are not covered in this thesis, it is worth noting that there are ongoing debates about the nature of human intelligence, the best way to measure intelligence and the inherent cultural biases of intelligence tests. Thus, there is controversy about the validity of any given measure of intelligence. Sparrow and Davis (2000) note that because intelligence is a psychological construct, the validity of any measure of it is largely dependent on the strength of the psychological theory on which the test is based.

& Plucker, 2013; McGrew, 2009). Within all CHC-models, there are dozens of narrow abilities, which are grouped beneath 8-16 broad abilities. There are differing opinions as to whether these broad abilities reflect an underlying 'general intelligence' with Carroll (1993, 1997) advocating in favour of general intelligence and Horn and Noll (1997) against.

Though the debate on the exact nature of intelligence is ongoing, a generalized measure of cognitive ability is widely accepted for use in longitudinal studies of development. More specifically, longitudinal studies often include measures of cognitive ability based on modern tests of intelligence developed using *CHC-theory*. These intelligence tests tend to provide a generalised score of intelligence (cognitive ability) along with subscores for each of the different categories of cognitive ability (Schneider & McGrew, 2012). Commonly used tests of intelligence include the Weschler Intelligence Scale for Children (WISC) and the Kaufman Assessment Battery for Children (K-ABC), both of which have been updated in the last two decades to capture features of CHC-theory (Alfonso, Flanagan, & Radwan, 2005). The K-ABC and WISC are both designed to capture several of the broad cognitive abilities included within CHC-theory and then synthesize them into a single cognitive score for use by practitioners and researchers. While the existence of a single general intelligence is still debated, using a combined measure is a common approach in various strands of research (e.g. Mistry et al., 2008; Spira & Fischel, 2005; Todd & Wolpin, 2003). The construct and criterion validity of these measures is widely explored in the literature. Thus, the results from any empirical study using these measures can be easily compared with existing research.

Cognitive Ability within this Thesis: In Section 2.1, cognitive ability was defined as an aggregate measure of a broad set of cognitive skills. For the empirical chapters of this thesis, the specific cognitive skills contained in this aggregate measure will be largely driven by the types of cognitive assessments available in the data. Fortunately, as mentioned above, the cognitive assessments used in longitudinal studies are typically taken from existing measures of general intelligence. More specifically, longitudinal surveys often include measures that closely align with the broad cognitive traits described by Carroll (1993). The exact choice of broad abilities used in my models will vary between empirical applications and further details are provided in each of the empirical chapters that follow. Using the construct of general intelligence to define cognitive ability not only aligns with CHC-Theory, but also closely corresponds with how economists define cognitive ability within human capital literature.[27]

[27]In the economics literature, the origins of cognitive measures is rarely discussed, a general cognitive ability is assumed, and the debate about the nature of intelligence is largely avoided.

Non-Cognitive Skills

Though the extensive literature discussed in Section 2.2 confirms that a wide variety of non-cognitive traits and abilities are robust predictors of adult outcomes, the majority of this research examines specific traits and abilities in isolation. As discussed in Section 2.3, economists have recently begun grouping together personality traits, along with distinct social, behavioural and emotional skills under a single factor of *non-cognitive ability*. Just as a general measure of cognitive ability is used to capture the broad set of cognitive skills described by CHC-Theory, this growing body of economics literature assumes that a unidimensional measure of non-cognitive ability can be created by aggregating a set of behaviours, personality traits, and social skills.[28,29]

The existing human capital literature only provides a loose definition of the specific traits that fall under this broad "non-cognitive" ability, with this definition varying according to the specific study and available data. In a survey of existing work, Heckman and Kautz (2012) acknowledge that "these attributes go by many names in the literature, including soft skills, personality traits, ..., character, and socioemotional skills" (p.452). Similarly, Cunha et al. (2010) describe non-cognitive abilities as "personality, social and emotional traits"(p.884), and in their empirical analysis non-cognitive skills are measured with a questionnaire originally designed to capture childhood psychopathology.

While there is substantial variability in the specific skills which form the aggregate measures of non-cognitive ability, many researchers include measures from psychometric instruments that were originally designed to capture the psychological constructs of *personality* and *emotional intelligence*. Understanding these constructs provides valuable insight when empirically modelling skill formation as well as a theoretical foundation for the interpretation of any empirical findings. Each of these constructs is examined below.

Personality: Alongside the evolution of the modern construct of cognitive ability, psychologists spent much of the twentieth century forming a framework that can be used to explain the similarities and differences of personality between individuals. Using factor analytic measures, personality researchers aimed to isolate a parsimonious set of factors that could describe the individual differences observed in society (Winter & Barenbaum, 1999). Beginning in the 1970s, researchers began to adopt a widely accepted taxonomy of personality traits being categorized under five broad factors that Goldberg (1971) labelled the "Big Five"(McCrae & John, 1992). Personality models that use the Big Five are referred to as *Five-Factor Models*.

[28]For example, Heckman, Stixrud, and Urzua (2006) "show that a model with one latent cognitive skill and one latent non-cognitive skill explains a large array of diverse behaviors".

[29]Empirical studies by Cunha and Heckman (2007, 2008) and Cunha et al. (2010) all use factor analysis to predict single latent factor scores to measure non-cognitive ability.

The details and specific names of the Big Five have evolved during the past 50 years, to the modern Big Five Factors with the acronym *OCEAN*. The five factors included in OCEAN are: **O**penness to Experience; **C**onscientiousness; **E**xtraversion; **A**greeableness; and **N**euroticism.[30] Though the Big Five were not originally designed to measure 'non-cognitive skills', it is possible to see how each of these dimensions could describe the *soft-skills* that would translate to academic and labour market success. While a full exploration of the history of personality psychology is not necessary for the present study, several key details of the Five-Factor-Model provide useful context for understanding its ability to capture non-cognitive ability.[31]

Under the Five-Factor-Model, each factor captures a set of more narrowly defined traits. These traits can be measured using self-report questionnaires, or through observation of an individual's behaviour. Proponents of the Big-Five argue that almost all measures of human personality can be mapped onto one of the five factors (Costa & McCrae, 1992; Goldberg, 1993; John, 1990; McCrae & John, 1992). Opponents of the model propose different factor structures to explain the set of human personality traits, though generally agree upon the underlying measures (John & Srivastava, 1999).[32]

Although there is ongoing debate about the factor structure of personality and the ability of the Five-Factor-Model to fully capture human personality, the framework of the literature "provide[s] a common language for psychologists from different traditions, a basic phenomenon for personality theorists to explain, a natural framework for organizing research, and a guide to the comprehensive assessment of individuals" (p.177 McCrae & John, 1992). In the context of this thesis, longitudinal data sets often make use of self-reported personality inventories. The questions contained in these inventories can be included in the aggregate measure of non-cognitive ability. Given the wealth of existing research that uses these personality inventories, it is possible to understand how these traits are distributed in various populations.

Emotional Intelligence: Although *emotional intelligence* has entered the general lexicon, it is rooted in psychological literature on human intelligence. Salovey and Mayer (1990) define emotional intelligence as "the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions"(p.189). Just as cognitive intelligence allows an individual to perform a variety of tasks, emotional intelligence plays a role in many daily interactions. In the same way that CHC-Theory informs the available measures of cognitive ability, theories of emotional intelligence have led to psychometric measures

[30] See McCrae and John (1992) and Costa and McCrae (1992).

[31] For the history of the Five-Factor-Model, the reader is directed to John and Srivastava (1999).

[32] For example, DeYoung, Carey, Krueger, and Ross (2016) and Hofstee, de Raad, and Goldberg (1992) present multi-dimensional extensions of the Big-Five, while Eysenck (1991, 1994) proposes a three-factor model which identifies: Psychoticism, Extraversion and Neuroticism.

of emotional intelligence that are often included in longitudinal studies. Though the research discussed in Section 2.2, provides clear evidence that these measures of *emotional intelligence* are robust predictors of academic and developmental outcomes, understanding how emotional intelligence is defined within the psychological literature provides insight into how these measures within my model of skill formation.

Although emotional intelligence was first mentioned in early models of general intelligence, an increasing interest in emotional intelligence beginning in the 1990s led to two modern conceptualizations of emotional intelligence: ability models and trait (or mixed) models.[33] Both of these models assume that emotional intelligence can change over time, but ability models focus on specific skills and abilities while trait models allow for more general measures of disposition. Extensive theoretical and empirical research exists for both types of models of emotional intelligence. This literature proposes specific instruments to measure each respective model of emotional intelligence, as well as extensive research measuring the validity of these instruments and how they relate to various academic and life outcomes. Although the full exploration of these different instruments is not relevant for this thesis, a basic understanding of each construct is useful as items from these inventories may be included in longitudinal data and can be used as part of an aggregate measure of non-cognitive ability.

Ability Models of Emotional Intelligence: Proponents of ability models argue that emotional intelligence refers to a discrete set of emotional skills and abilities. Mayer and Salovey (1997) proposed an ability model of emotional intelligence within which emotional skills fall under four branches: the ability to perceive and express emotion; the ability to assimilate emotion into thinking; the ability to understand and analyse emotion; and the ability to regulate emotion. Emotional intelligence is the combination of these branches and "can be assessed most directly by asking a person to solve emotional problems, such as identifying the emotion in a story or painting, and then evaluating the person's answer against criteria of accuracy" (p.268 Mayer, Caruso, & Salovey, 1999). Using this logic, Mayer et al. (1999) developed the *Multifactor Emotional Intelligence Test* (MEIS) which assesses performance on a series of emotional tasks.[34]

Trait (Mixed) Models of Emotional Intelligence: While ability models of intelligence only include measures of performance on specific tasks, mixed models of emotional intelligence "encompasses behavioural dispositions and self-perceived abilities

[33]For a detailed discussion of the major theoretical models of emotional intelligence the reader is directed to R. D. Roberts, Schulze, and MacCann (2008).

[34]Improvement of the original MEIS resulted in its successor the *Mayer-Salovey-Caruso Emotional Intelligence Scale* (MSCEIT Mayer, 2002).The MSCEIT assesses performance on a variety of emotional tasks using a 141 item test. Since it was introduced, the validity of the MSCEIT, and how it relates to various measures of success has been assessed in a variety of samples (see for example: Mayer, 2008; Mayer, Caruso, & Salovey, 2016; Mayer, Panter, & Caruso, 2012; Mayer, Roberts, & Barsade, 2008; Mayer, Salovey, Caruso, & Sitarenios, 2003).

and is measured through self-report"(p.426 Petrides & Furnham, 2001). Several trait models of emotional intelligence have been proposed[35] and all such models suggest that emotional intelligence is best measured using self-report questionnaires. The two most widely used questionnaires are the *Trait Emotional Intelligence Questionnaire* (TEIQue; Petrides & Furnham, 2003) and the *Emotional Quotient Inventory* (EQ-i; Bar-On, 1997).

Using Measures of Emotional Intelligence: Whether defined using measures of performance on specific tasks (ability based models), or using a self-reported assessment of an individual's characteristics (trait based models), there is extensive research demonstrating the predictive power emotional intelligence inventories. In the context of this thesis, both self-report and task-based assessments of emotional intelligence can be considered as possible measures when defining non-cognitive ability. The final choice of measures will be driven by what types of assessments are included within the data used for each empirical application. As with personality traits, all measures of emotional intelligence should be assessed within the existing literature and interpreted using research on the psychometric properties of a given instrument.

Non-Cognitive Ability within this Thesis: While the constructs of personality and emotional intelligence capture many components of non-cognitive ability, including them in an empirical model requires measures which adequately capture each construct. Although inventories exist to measure personality and emotional intelligence, they are not necessarily adapted to children. Even when an age-appropriate inventory exists, it may not be included the available longitudinal data. Fortunately, many of the abilities; individual traits; and behaviours that are captured by personality and emotional intelligence are included in scales originally developed to measure childhood psychopathology; to measure child temperament; or to identify children that are falling below developmental milestones. Longitudinal surveys often include these types of scales, which can serve as suitable proxies for measuring the non-cognitive skills discussed above.

The specific scales used to capture non-cognitive ability will depend on the available data. There is a growing body of research linking existing survey instruments to the measures of emotional intelligence and personality introduced earlier in this section.[36] For the empirical applications included in this thesis, each specific scale and how it relates to personality and emotional intelligence will be evaluated on a case-by-case basis.

[35]Examples include: the Bar-On Model of Emotional-Social Intelligence (Bar-On, 1997, 2004, 2006, 2010); Goleman's Emotional and Social Competence Theory (Goleman, 1995) and Petrides and Furnham's model of emotional intelligence as a distinct personality trait (Petrides & Furnham, 2001; Petrides, Pita, & Kokkinaki, 2007).

[36]Specific examples include: the work of Almlund et al. (2011) who discuss how measures of child temperament correspond with the Big Five personality traits; the studies of Mavroveli, Petrides, Shove, and Whitehead (2008) and Poulou (2014) which both find a correlation between teacher reported child development (using the Strength and Difficulties Questionnaire) and trait emotional intelligence scores measuring the TEIQue; and the research of DeYoung et al. (2016) that finds a relationship between the Big Five personality traits and measures of psychopathology included in the DSM-5.

2.4.2 Theoretical Models of Child Development

There is no consensus amongst psychologists regarding a unifying model of child development. This is because "no single theory captures all of the complexities of human development" (Cook & Cook, 2013, p.18). Instead, contemporary understanding of child development relies on a range of developmental theories. By combining key aspects of these theories, researchers create a theoretical foundation which can be used to explain different aspects and stages of a child's development (Santrock, 2017).

This theoretical eclecticism acknowledges specific criticisms of each model and proposes drawing on multiple models to explain child development (Shaffer, 2014). This thesis draws on six developmental theories to capture various aspects of development. These theories are:

- Piaget's Cognitive Developmental Theory,
- Vygotsky's Sociocultural Theory,
- Information-Processing Theory,
- Erikson's Psychosocial Theory,
- Ethological Theories of Development (including Bowlby's Attachment Theory),
- Bronfenbrenner's Ecological Systems Theory.

Contemporary developmental psychology research rarely uses these theories as the sole explanation for development but instead draws on aspects of the models to provide a general framework for developing specific theories about one aspect of development. These modern theories then inform research on how children develop in a specific domain (i.e. language acquisition or working memory). Therefore, the six theories listed should not be thought of as fully defining skill development, but instead as providing the tools with which to identify specific inputs and to evaluate an overall model of skill formation.

Although psychologists will be acquainted with these six theories of child development, a concise summary of the models, along with their relevance for this thesis, is presented below for researchers from other disciplines who are unfamiliar with these concepts.[37] As discussed above, no single theory can explain all of human development and contemporary theories often apply to the development of a specific skill. These models just provide a general framework for understanding the broader developmental trajectories. The empirical model draws on features of each model as outlined below, while acknowledging that another model may better explain other areas of development. These models alone are not sufficient to explain development, but instead provide different ways of conceptualising skill formation and help identify potential measures to include in the model.

[37] A full review of developmental psychology is beyond the purview of this thesis. The introductory textbooks of Keenan, Evans, and Crowley (2016), Santrock (2017), and Meadows (2017) review the major developmental models and their contribution to contemporary understanding of child development.

Piaget's Cognitive Developmental Theory: Piaget (1952a, 1952b, 1954) proposed that a child's development follows a universal sequence of stages in which children develop cognitive *schemas* in order to understand the world around them. In each of these stages, children *assimilate* knowledge that is consistent with their existing understanding of the world and adjust their schema in order to *accommodate* new knowledge that conflicts with their previous understanding. Piaget proposed that each stage is distinct, and children have qualitatively different ways of thinking in each stage. A child's developmental stage determines how they will respond to various influences from the environment. Progression through these stages is a natural process whose speed is based on the innate traits of each child. Thus, the development of new skills is determined by the child's psychological maturation and their existing social and cognitive capacities.

Piaget believed that children's progression through these distinct stages resulted in shifts in thinking across all areas of development and that this pattern of development is the joint result of interaction with the environment and biological maturation. Although more recent research has shown that children do not follow the strict developmental stages described by Piaget, modern developmentalists do agree that children's thought patterns change over time and this development results from a combination of biology and the child's interaction with the world around them (Harris & Westermann, 2014).

Piaget's theory implies that a given parental input is only effective once the child reaches the developmental stage where they possess the skills to assimilate the parental stimulus into their understanding of the world. Consequently, empirical models of skill formation should allow the impact of parental inputs to change over time and account for the interaction of previous skills and environmental inputs in the formation of new skills.

Vygotsky's Sociocultural Cognitive Theory: Like Piaget, Vygotsky (1962) believed that biological and environmental factors interact to shape a child's development. However, while Piaget proposed a progression through universal developmental stages, Vygotsky believed that the developmental process varies between individuals and progression is largely driven by the social and cultural context in which the child is raised (Shaffer, 2014). In his work, Vygotsky proposes that social interactions allow children to master culturally specific psychological tools such as language. Using these tools, the child develops an understanding of the world around them. Through continued social interaction the child develops further culturally specific tools which leads to their world understanding evolving over the course of childhood (Santrock, 2017).

With regard to models of skill formation, Vygotsky's Sociocultural Model supports the inclusion of measures of parental input across a variety of activities and for the specific inputs to be driven by the cultural context of the children in question. While the Piagetian model also allows for diverse inputs, it is more focused on the timing of these inputs and how they promote the development of an existing sequence of skills.

Information-Processing Theory: According to information-processing theory, development results from the gradual improvement in the child's ability to: process and store information from their environment, retrieve information from their memory, and to use this information in order to master increasingly advanced skills (Santrock, 2017). Using computers as an analogy, information-processing theories posit that through the process of biological maturation, a child's mind develops more advanced 'hardware' in the form of higher innate capacity for processing information. Similarly, interaction with the environment builds the knowledge and skills that form the 'software' of human cognition.

Information-processing theories support the need for skill-formation models that allow for inputs to occur at different points in physical development. Similarly, models must allow for previously demonstrated skills to interact with current investment to shape future abilities. In the context of information-processing, environmental inputs will not shape the child's development if the existing skills do not allow the child to interpret and store the information from their environment.

Erikson's Psychosocial Theory: Erikson (1950, 1968) proposed a stage-theory which focuses on the development of a child's personality. In Erikson's model, a child progresses through eight stages of development where the driving force is a quest for identity. These stages occur in a pre-specified sequence and their timing is the joint result of biological maturation and the child's experience of his or her environment (Shaffer, 2014). During each stage, the individual faces a psychosocial conflict and the resolution of each conflict is needed for further development during the next psychosocial stage.

Erikson's Psychosocial Theory has several implications for modelling skill formation. Without exploring the specifics of each stage, it is clear that Erikson's theory provides support that personality development at a given stage requires not only environmental inputs, but also pre-existing skills. This corresponds with an empirical model that allows for cross-productivity of skills as well as time specific measures of the return to parental investments. Just as Piaget's stages of cognitive development indicate that parental investment is only effective in promoting cognitive development in the presence of certain skills, Erikson's stages of psychosocial development indicate that non-cognitive development is the interplay of both prior skill and environmental inputs.

Ethological Theories of Development: Ethological theories of development propose that development follows a biologically programmed path whereby the child instinctively responds to the environment in the way that is most likely to guarantee his or her survival. If a certain environmental stimulus is absent from a specific period in the child's life, these instinctive responses will not occur, and the child will fail to develop the corresponding skills. Ethologists propose that there are *critical periods* for development, which are "a limited time span during which developing organisms [children] are biologically prepared to display adaptive patterns of development, provided they receive appropriate input" Shaffer (2014, p.60). The environmental input is only effective at promoting development during this critical period and will have no effect if it occurs at another point in the child's development. Similarly, during *sensitive periods* the child is most responsive to a certain input, and investments made during other stages will be less effective.

One particularly prominent ethological theory is Bowlby's Theory of Attachment (Bowlby, 1969), which proposes that children who are not provided with the correct environmental stimulus in infancy will fail to form appropriate attachments with their caregivers. According to Bowlby, failing to establish proper attachments makes it difficult for individuals to develop appropriate social and emotional skills later in life.

The evidence presented by ethologists indicates the presence of critical and sensitive periods of development is central to the model of skill formation presented in this thesis. In order to capture these periods, the model must allow for age-specific returns on parental inputs, along with the requirement for future skills to build on prior skills. These concepts will be addressed in greater detail in Chapter 3.

Bronfenbrenner's Ecological Systems Theory: Ecological systems theories propose that development is the result of interactions between the child and various aspects of his or her environment. The most prominent ecological systems theory was introduced by Bronfenbrenner (1979) and further expanded by Bronfenbrenner and Morris (1998, 2006). According to Bronfenbrenner, a child's environment consists of multiple contexts, the confluence of which shapes the child's developmental trajectory. Not only does the child's development depend on the interaction of multiple inputs, but Bronfenbrenner and Morris (2006) specifies that "[t]o be effective, the interaction must occur on a fairly regular basis over extended periods of time."(p.797).

For this thesis, Ecological Systems Theory implies that models of skill formation would ideally include measures of various features of the child's environment such as siblings, socio-economic conditions, religious interactions, etc. Additionally, Ecological Systems Theory would suggest that any measurements of parental input must examine consistent patterns of behaviour, and that the model must account for the evolving interaction between the child and his or her environment.

2.4.3 Theoretical Models of Parenting

The empirical research discussed in Section 2.2 highlights the long-standing interest in measuring how parental behaviours, attitudes and expectations shape children's development. In addition to this empirical work, there is an extensive literature which presents theoretical models of how parent-child-interactions drive skill formation. Although models of child development and theoretical models of parenting are presented separately within this thesis, it is important to acknowledge there is substantial overlap between the two sets of theories. Specifically, many of the models of development that have already been presented in Section 2.4 define precise roles for parents within their framework for child development. For example, in the socio-cultural model of development introduced in the previous section, Vygotsky (1978) discusses how parental behaviours help to guide development. The previous subsection has already addressed the role of parenting in each of the theories of development, but alongside the theories of child development, there are many stand-alone theories to define the nature of parent-child-interaction.

To review the contemporary models of parenting I follow the terminology presented in the *Contextual Model of Parenting* introduced by Darling and Steinberg (1993). They propose that in order to understand the influence of parenting on child development, researchers must distinguish between three aspects of parenting: parenting goals, parenting practices and parenting style. The parenting process begins with *parenting goals*, which are "the values parents hold and the goals toward which they socialize their children" (Darling & Steinberg, 1993, p.492). These goals directly influence: the specific behaviours which form the *parenting practices* that parents employ, and the *parenting style* which describes the parental environment in which these behaviours are practised. More specifically: Darling and Steinberg (1993) define "parenting style as a constellation of attitudes toward the child that are communicated to the child and that, taken together, create an emotional climate in which the parent's behaviors are expressed" (p.488).

The role that the three aspects of parenting play in the Contextual Model of Parenting are illustrated in Figure 2.1. According to Darling and Steinberg (1993), parental goals and values do not directly impact children's development, but instead influence parenting style and parenting practices. Both parenting style and practices then go on to influence a child's development. Within the model, "parenting practices have a direct effect on the development of specific child behaviours" while "the primary processes through which parenting style influences child development are indirect" as "parenting style can best be thought of as a contextual variable which moderates the relationship between specific parenting practices and specific developmental outcomes" (Darling & Steinberg, 1993, p.493). If parental goals and values only impact development via parenting style and practices, then it is possible to model the direct effect of specific parenting practices and style without measuring the motivation behind these inputs. This implication from the Contextual Model of Parenting is consistent with how economists conceive of parental motivations and provides further support for the validity of modelling the effect of specific parental inputs while ignoring the motivation behind these inputs.

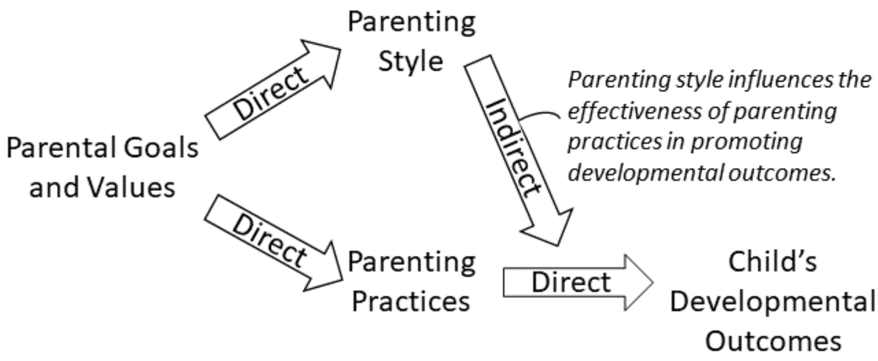


Fig. 2.1 Contextual Model of Parenting: Adapted from *Figure 1*, Darling and Steinberg (p.493:1993)

Although the terminology presented by Darling and Steinberg's Contextual Model of Parenting is useful to frame the discussion of parenting, and the terms *parenting practices* and *parenting style* are widely cited by the literature, the mechanisms presented by Darling and Steinberg (1993) have yet to be proven. More specifically, while there is evidence that parental goals and values directly shape parenting style and practices (examples include: Bradley & Corwyn, 2002; Brooks-Gunn & Markman, 2005; Whiteside-Mansell, Bradley, Tresch Owen, Randolph, & Cauce, 2003) the specific pathways that Darling and Steinberg (1993) propose between parental practices, parental style and child outcomes are not fully supported by empirical research. Spera (2005) argues that although there is strong evidence linking parenting styles and developmental outcomes, "little research to date has examined whether parenting styles actually moderate the relationship between parenting practices and adolescent outcomes" (Spera, 2005, p.140).

As there is substantial evidence that both constructs influence child development, I propose empirical models include measures of both parental practices and style. From a policy perspective, parenting style and practices are more likely to respond to public policy interventions compared to parental values and goals. Unlike parental goals and values, which are largely based on personal preferences, the behaviours that constitute parental practices and style are displayed across a range of individuals and cultures.

While my empirical modelling strategy does not employ the pathways proposed by Darling and Steinberg (1993), their terminology is widely used. For this reason, it is important to understand how the terms *parenting style* and *parenting practices* are used by psychologists. These terms have been used in a variety of parenting models with researchers studying how these behaviours apply in different types of families (e.g. single parents), in different cultures and in non-parental caregiving situations. An understanding of these constructs allows researchers to accurately identify the measures of parental input to include in any empirical model of skill formation. Below, I define parenting style and parenting practices and discuss the implication of these constructs for my research.

Parenting Style: The modern construct of *parenting style* stems from the seminal works of Baumrind (1966, 1967, 1971), which are based on the belief that "it is more meaningful to talk about the effects of patterns of parental authority than about the effects of single parental variables"(Baumrind, 1971, p.95).[38] Using the findings from multiple observational studies of the interactions between parents and their children, Baumrind (1966) includes measures of multiple aspects of child-rearing behaviour such as discipline, encouragement, communication, nurturing, and involvement to identify three parenting typologies: authoritative, authoritarian, and permissive. A full description of these typologies is provided by Baumrind (1966), but for the sake of brevity, I direct the reader to the succinct definitions provided by Baumrind (1971):

- *Authoritative* parents are "controlling and demanding [towards their children]; but they were also warm, rational, and receptive to the child's communication"(p.1).
- *Authoritarian* parents are "detached and controlling, and somewhat less warm than other parents" (p.2).
- *Permissive* parents are "noncontrolling, nondemanding, and relatively warm."(p.2).

Building on these three typologies, Maccoby and Martin (1983) introduced a two-dimensional framework which defined parenting styles as the intersection of responsiveness (support, warmth, involvement, acceptance) and demandingness (control, restrictions). Under this framework, a fourth parenting typology emerged: *neglectful* parenting. The two-dimensional framework defines the four types of parenting styles (authoritative, authoritarian, indulgent, and neglectful) based on the level of responsiveness and demandingness displayed by parents. I have illustrated this typology in Figure 2.2.

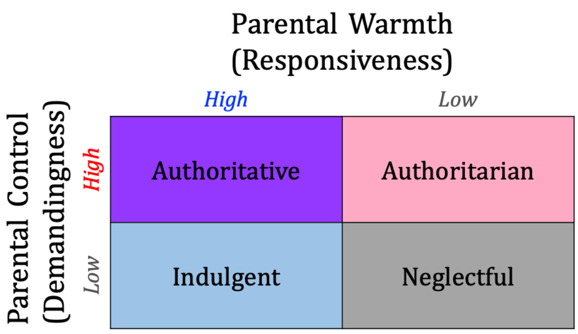


Fig. 2.2 Parenting Style Typologies

Further work by Baumrind (1991, 2005) confirms these two dimensions of parenting and provided evidence of the relationship between parenting style and various aspects of child development. The four parental styles have been repeatedly used to predict a variety of developmental outcomes with authoritative style parenting generally corresponding with higher levels of social, emotional and cognitive development (examples include:

[38] Baumrind uses the term "variables" to describe individual parental behaviours. Instead of measuring the specific impact of these individual behaviours, Baumrind proposed focussing on larger constructs.

Areepattamannil, 2010; Lamborn, Mounts, Steinberg, & Dornbusch, 1991; Lee, Daniels, & Kissinger, 2006; Steinberg, Lamborn, Dornbusch, & Darling, 1992).[39]

When identifying parental inputs for empirical models of skill formation, possible measures can include the traits and behaviours provided by Baumrind (2005) in her detailed descriptions of the dimensions of responsiveness and demandingness. According to Baumrind (2005): *responsiveness* is "the extent to which parents foster individuality and self-assertion by being attuned, supportive, and acquiescent to children's requests; it includes warmth, autonomy support, and reasoned communication" (p. 61). The dimension of *demandingness* describes "the claims parents make on children to become integrated into society by behavior regulation, direct confrontation, and maturity demands (behavioral control), and supervision of children's activities (monitoring)." (Baumrind 2005, p. 62). Any measures that capture these various aspects of parenting can be thought to contribute to child development and help provide a valuable framework for defining parental investment through the lens of parenting styles.

Parenting Practices: While parenting style refers to general patterns of parent-child interaction, parenting practices are defined as "specific, goal-directed behaviors through which parents perform their parental duties" (Darling & Steinberg, 1993, p.488). Parenting practices as defined by Darling and Steinberg (1993) are what other authors might refer to as parenting behaviours. Unlike parenting styles which require measuring several dimensions of child-rearing, parenting practices can be measured in isolation. As seen in Section 2.2, there is a wealth of research linking specific parental behaviours to a range of developmental outcomes. Although recent psychological studies of parenting have focused on the relationship between parenting style typologies and developmental outcomes, there remains substantial research to be done to identify the role of specific parenting behaviours. A focus on parenting practices is especially relevant when examining specific aspects of development with Carlo, McGinley, Hayes, Batenhorst, and Wilkinson (2007) noting that "there is growing interest in identifying specific parenting practices, rather than assessing general interaction styles to better predict [specific] behaviours" (p.147).

From a modelling perspective, parenting practices tend to vary more over the life of a child, which allows for a more nuanced understanding of the role that parenting plays at specific stages of development. As part of a longitudinal study of parenting behaviour and child development, Gutman and Feinstein (2010) note that "[a]lthough there is considerable continuity in parents' child-rearing orientations [style], parents modify their behaviours in response to their children's developing abilities and needs" (p.536). This dynamic nature of parenting practices suggest that they might provide more meaningful predictions of cognitive and non-cognitive development within a longitudinal model.

[39] These studies were largely conducted using children from American middle-class households. There is research to suggest that these patterns vary across national and cultural contexts.

Parental Investment within this Thesis

Although the review of empirical literature showed the wide range of parenting behaviours, child-rearing strategies, parenting styles and family goals that have been used to predict childhood developmental outcomes, the models of parenting discussed previously in this chapter show how these measures fit within theoretical constructs used to define parenting. For this thesis, this theoretical understanding of parenting provides an important framework and relevant terminology for specifying parental investment and adds to the understanding of investment provided by the economic models of skill formation.

While the exact specification of parental investment will vary depending on the available data, there is substantial evidence that variables measuring parenting style and parenting practices should both be considered possible measures of parental input. Based on the framework presented by Darling and Steinberg (1993) measures of specific parenting practices will more directly influence children's development, but the works of Baumrind and others show that parenting style is still an important consideration. Regardless of the specific measures chosen, any interpretation of results should use the theoretical models of parenting described above to contextualize the findings and explain the pathways through which parenting is impacting children's development.

2.5 Using Theory: Education Research

The two previous sections in this chapter introduced the relevant theoretical frameworks used by psychologists and economists to explain skill formation in children. Although I have chosen to discuss them within the context of their respective broad disciplines, in both economics and psychology the relevant theories are derived from an interdisciplinary history of research which aims to understand the process of learning. When investigating how children learn, researchers in psychology and economics draw on certain aspects of the other discipline but tend to focus on the theoretical and empirical approaches from their respective fields.

It is under the umbrella of education that these two disciplines have historically come together. Across all fields of study, research on skill development has frequently been motivated by the intention of measuring the efficacy of formal education. For example, the human capital models of skill formation discussed in Section 2.3 were originally motivated by a desire to explain the role that education, as a major form of investment in skill development, plays in determining wages in the labour market. As mentioned in Section 2.4 one of the original motivations behind the development of intelligence tests was to provide a scale to be used by education systems. Similarly, there is substantial crossover between the theoretical frameworks child development discussed in Section 2.4 and the philosophical debates about the nature of the education system.

Due to the interdisciplinary nature of their research, educationalists draw on both economics and psychology to model child development. The field of education creates the space where these two theoretical perspectives can be combined to form a new framework to explain child development. As the modelling considerations derived from each theoretical perspective have already been discussed in Section 2.3 and Section 2.4, these theories are not reintroduced within the context of education. Instead, this section focuses on the distinct contribution that educationalists make to understanding skill development through the education literature focused on modelling school readiness.

As presented in Section 2.2, many prior studies of childhood ability were designed to assess a child's readiness for formal education. By reviewing the theoretical frameworks surrounding school readiness, I am able to show not only how this literature informs longitudinal studies of skill development, but also how the conceptual model of school transition used by educationalists could extend to cover the full span of development.

Though the school readiness literature typifies how educationalists draw on theory and evidence from multiple disciplines, it is by no means the only strand of research in the education literature where skill development is examined. Instead, it is one example of why the type of multidisciplinary approach proposed by this thesis requires both the empirical evidence and the integrative approach provided by the field of education.

School Readiness

The ability to predict a child's achievement in the early years of formal education is a well-studied topic. This research is motivated by the assumption that early academic success is a stepping-stone for a successful path through formal education, which will in turn yield positive outcomes in adulthood. Research in this field aims to identify the behaviours, skills, and traits of the child, as well as the factors in the child's environment that may predict a successful transition to formal education. Once identified, these predictors of success are used by educators to identify children at risk of being ill-prepared for formal education and falling behind their peers.[40] The findings of this research are also used to create intervention programs to mitigate the identified risk factors.

The major theoretical framework used to explain school readiness conceptualizes a child's entry to formal education as a developmental period of transition that is influenced by multiple factors and spans several years (Petriwskyj, Thorpe, & Tayler, 2005). McClelland et al. (2000) note that a successful transition to school lays the groundwork for future academic success and that "learning-related skills [acquired through this transition] continue to be linked to a child's academic success... and provide the foundation for later academic performance" (p. 492). Educators and policy-makers are especially concerned with a successful transition to school, and children are often assessed on their early progression in their classes.

Although these school transition models are focused on one period of childhood transition and have not been extended beyond this period, they provide valuable insight into the way educationalists model development. They also capture much of the joint role that cognitive and non-cognitive skills play in successfully navigating this period of a child's life. Vogler, Crivello, and Woodhead (2008) provide a detailed discussion of school transition literature from the last century and explain how theoretical understanding has shifted from the early child-centered models of development, towards socio-cultural models, and finally to models which combine aspects of both child-centered and socio-cultural theories. A brief review of these models helps to illustrate the substantial overlap between this education research and the models introduced in the previous sections. Additionally, the findings obtained from research on school readiness that is based on these models can inform the decisions made about how to model skill development

Child Centered Models

Child centered models are strongly influenced by the theories of Piaget and Erickson. As discussed in Section 2.4, Piaget and Erickson believed that a child's development follows a sequence of stages in which children develop and transform their cognitive, emotional, physical, social, and psychological skills Vogler et al. (2008). Stage-theory

[40]For this reason, research in the area is sometimes referred to as *assessing school readiness*.

argues that progression through these stages is a natural process whose speed is based on the innate traits of each child. Thus, the transition to school, and the corresponding development of new skills, is determined by the child's psychological maturation and the development of their social and cognitive capacities. In this context, school readiness is the point when a child has reached the right developmental stage to have the proper skills to function in the educational environment. Prior to this stage, parental investment into skill development is needed to ensure that the child has the proper skills to function in the educational environment. This is referred to as *scaffolding*.

Socio-Cultural Models

On the other hand, socio-cultural models argue that the development process is driven by the inputs from the child's home and school environment. Socio-cultural models build on the work of Vygotsky who, as discussed in Section 2.4 emphasized the role that the child's environment plays in development. Specifically, how the interaction between the child and their environment can change the way in which the child responds to certain stimulus and develops skills.

Combined Models:

Similar to the theoretical eclecticism of developmental psychologists, educationalists have recently proposed that neither the child-centered, nor the socio-cultural models fully capture the transition to school. This has led to the creation of *combined models of school transition.* The most representative model is the The Ecological and Dynamic Model of Transition (EDMT) presented by Rimm-Kaufman and Pianta (2000). This model posits that academic success is a dynamic process, resulting from "the combined influences of child, direct, indirect, and dynamic effects of contexts on children's transition to kindergarten" (p. 499). In essence, this model is a synthesis of the child-centered and socio-cultural models. As well as including features of both types of models, the EDMT stresses that development is a dynamic process in which the original inputs interact to determine the inputs in the next period.

Within Rimm-Kaufman and Pianta's model, the traits and skills of a child at the beginning of school shape the way that the child interacts with the resources presented to them through formal education. The interaction between these pre-existing factors determines the rate at which the child is able to acquire new skills. As the child develops, their new skills shape the way in which they interact with the various factors in their environment, sometimes even resulting in changes in the environment itself. In turn, the child is exposed to new inputs, which shape the child's development going forwards. For example, as a child develops relationships with their teachers, the child might act in such a way that prompts the teacher to provide educational interactions that directly results from the way the child behaves. The continued interaction between the child and their environment is what shapes the child's transition to school. Furthermore, "these interactions, over time, form patterns and relationships that can be described not only as influences on children's development, but also as outcomes in their own right" (Rimm-Kaufman & Pianta, 2000, p.499).

How does this School Readiness Model Inform Research on Skill Formation?

Though the EDMT was designed to explain the transition to kindergarten, Rimm-Kaufman and Pianta (2000) frequently refer to its ability to capture the development of skills. For this reason, the EDMT is a good example of how one specific model builds upon several theories to apply them to a specific education context.

When examining the EDMT with the economic model of skill formation in mind, it is easy to see how this theoretical model has many similarities to the empirical model presented by Cunha and Heckman (2007). For example, throughout their discussion of the EDMT, Rimm-Kaufman and Pianta (2000) emphasize that skill acquisition is a dynamic process, such that each successive level of achievement and skill development is determined by the interaction between new inputs and stimulus from the child's environment and skills and traits developed in prior periods. More simply put, there are numerous characteristics and environmental factors that can interact to either directly produce academic achievement, or to produce other skills which in turn lead to future achievement. This supports many features of the skill formation model presented in Section 2.3, along with providing valuable insight into the types of input and outcomes to include in any empirical work.

2.6 Implications for this Thesis

The sections above have presented a broad overview of the existing empirical work on parenting and skill development as well as detailed examinations of the most pertinent theoretical frameworks from economics, psychology and education. This overview has shown that each of these theoretical models provides a slightly different way to examine the role that parents play in their children's skill formation. I argue that none of these models fully capture the true nature of skill development and that by drawing on theoretical aspects of each strand of research I am best able to model skill development. Using this combination of underlying theory has major implications for the choice of statistical methodologies, as well as the conclusions that can be reached from my results.

I conclude this chapter with a brief review of the key considerations that can be taken from each discipline, as well as the general implications of this literature for empirically modelling skill development.

Relevant Considerations from Economics

The literature from economics not only provides the theoretical constructs used to describe parenting as an investment in human capital but also presents modelling strategies for empirically estimating such human capital frameworks. Section 2.3 presented a detailed mathematical representation of the production functions for skill formation that will form the basis of the theoretical framework presented in Chapter 3. This production function literature helps to conceptualize parenting as an investment, and provides many of the mathematical tools used to measure it.

It is clear from the human capital literature that there is a need to define models that capture the multidimensional nature of human ability. Similarly, recent works by Cobb-Clark, Nicolás Salamanca, and Zhu (2018); Del Boca et al. (2017); Bono et al. (2016); and Doepke and Zilibotti (2017) have all shown the more nuanced understanding that can be gained from defining parental inputs using multi-dimensional measures that do not only focus on household resources. Unfortunatelyf, there is limited research which simultaneously estimated the development of multiple types of ability while using multi-dimensional measures of parental inputs. By defining a model which captures both features, my research will address this gap in the literature.

Relevant Considerations from Psychology

While the model I present in this thesis is grounded in economics, literature from psychology helps refine the model specification and allows me to understand the constructs behind existing measures when selecting model inputs. More precisely, psychometric literature is crucial for defining the measures of both cognitive and non-cognitive ability in this thesis. Though these measures are widely accepted and often well validated, regardless of the theoretical framework used for an analysis, it is vital to understand the underlying basis for each measure included in the model.

In addition to helping measure the inputs to my model, psychological theories also inform the theoretical framework presented in Chapter 3. Though the underlying mathematical framework is rooted in economic models of human capital, the theories of child development from Section 2.4 provide valuable context for defining the relationships between various components of the model. Specifically, although the historical theories of child development have been replaced by more domain-specific theories, the general frameworks from these models, which contemporary developmental psychologists use to inform modern theories of development can help provide perspective for the way I specify my model. For example, the stages of development proposed by Piaget support empirical models with multiple periods of childhood, while socio-cultural models of development imply that family characteristics should be considered when defining parental inputs.

Finally, the psychological literature on parenting of Baumrind (1966, 1971, 1978), Maccoby and Martin (1983), and Darling and Steinberg (1993) provides strong justification for the importance of examining multiple types of parental input. This literature shows us that not only are there multiple ways to conceptualize parental input, but that both theory and data must be used when identifying specific parental inputs.

Relevant Considerations from Education

The literature from education provides many of the empirical findings regarding the role that a variety of skills play in determining academic success. I have shown in Section 2.5 how the theoretical framework behind school readiness literature captures many of the theoretical considerations contained in the economic and psychology models of skill formation. As these models are defined to capture development across childhood, it is reasonable to assume that the findings from education research extend beyond school transition. However, even if the empirical findings only apply to school transition, they still provide valuable information about what factors to include within a larger model.

Literature on transition to formal education also highlights the importance of modelling skill development as a dynamic relationship. Any analysis that simply categorizes various observable factors as ‘predictors’ and ‘outcomes’ may fail to capture the way that very early development prompts later aspects of childhood skill formation. Rimm-Kaufman and Pianta (2000) allude to this empirical consideration when they argue it is “critical that the interactions among the [child, their skills and their environment] be measured repeatedly and longitudinally” (p. 504).

Consolidating these Fields

Taken together, the literature on childhood skill formation and the role of parental inputs provides seven general findings that inform the empirical work of this thesis and ensure it makes a novel contribution to the existing literature on childhood skill development. These are each listed below.

Finding 1: *Childhood consists of multiple periods of development.* Theories of child development presented in Section 2.4 provide strong evidence that childhood consists of multiple stages and that it is unrealistic to model it as a single period.

Finding 2: *Skill development is cumulative.* The concept of *scaffolding* from developmental psychology, as presented in Section 2.4, as well as the economic evidence of self-productivity of skills, as presented in Section 2.3 both support the idea that skills created in one period continue to exist in the next period.

Finding 3: *Skills might develop differently at different stages of development.* The theoretical models presented in both Section 2.3 and Section 2.4 demonstrate that the efficacy of various investments differs depending on when the investment is made. To capture the changing nature of skill development, empirical work must allow for stage-specific estimates on the effect of parenting on skill development.

Finding 4: *Cognitive and non-cognitive skills cannot be examined in isolation.* The production functions discussed in Section 2.3 provide substantial evidence to support models which simultaneously estimate the formation of cognitive and non-cognitive skills. Similarly, the school-readiness literature discussed in Section 2.5 demonstrates the inaccuracy of modelling different types of ability in isolation.

Finding 5: *Parental investment is multidimensional.* There is extensive evidence from both psychology and economics to support the idea that there are multiple types of parental investment with distinctions made between financial resources, parenting behaviours, styles of parenting and family characteristics.

Finding 6: *Skills are self-productive.* The psychological theories presented in Section 2.4 and the economic models in Section 2.3 both provide theoretical and empirical support for the idea that skills acquired during one period of childhood foster further development of that type of skill in future periods.

Finding 7: *Skills can be cross-productive.* The human capital literature in Section 2.3 provides empirical evidence of cross-productivity. The school-readiness literature, as explored in Section 2.5, provides further support by highlighting the important role that certain behaviours and social skills play in the development of cognitive ability in primary school. Similarly, the psychology literature, examined in Section 2.4, shows that personality traits and emotional intelligence both foster cognitive development.

CHAPTER 3

Theoretical Framework

This chapter presents a theoretical framework and empirical model that can be used to measure the role of parenting in the development of cognitive and non-cognitive skills. As this framework draws on research from multiple fields, basic methodological details for each component of the model are included in the exposition of the model. While some readers will be well-versed in the statistical techniques examined below, and find these details overly simplistic, it is important to review each of the relevant estimation strategies for readers who are unfamiliar with any of the components of the larger model.

The theoretical framework and the relevant empirical considerations are presented in five parts. First, Section 3.1 discusses how the seven key findings from the child development literature, summarised at the end of Chapter 2, inform the human capital production function that this thesis uses to model skill development. In this section, I define the basic components and structure of the production function and explain why I have chosen to model skill formation in this way. Secondly, after introducing the general structure of the production function, Section 3.2 provides a detailed exposition of the full empirical model that can be used to estimate this theoretical framework. This section includes specific details, not only on how to define the production function for skill formation, but also how to estimate this model using empirical data. Next, Section 3.3 lays out an identification strategy that draws on both theoretical and statistical models to define the different types of parental investment. Then, Section 3.4 reviews the data requirements and methodological considerations that I make when empirically estimating the proposed model using longitudinal survey data. The chapter concludes with Section 3.5 which discusses the implications that the proposed methodology has for the empirical research of this thesis.

3.1 Defining the Conceptual Model

At the end of the previous chapter, Section 2.6 identified seven key findings from the literature on child development and parenting that shape the theoretical framework used in this thesis. These seven aspects of the literature are presented again below, along with the implications that they each have for my proposed estimation strategy.

Finding 1: *Childhood consists of multiple periods of development.* A multi-period model must be used, with measures of skill and ability at each time point.

Finding 2: *Skill development is cumulative.* Using a recursive model allows for skills in one period to be a function of prior skills and the entire history of parental inputs.

Finding 3: *Skills might develop differently at different stages of development.* A suitable model must allow for stage-specific estimates. These can be used to identify critical and sensitive periods of investment.

Finding 4: *Cognitive and non-cognitive skills cannot be examined in isolation.* Skill must be modelled as a vector. This allows the model to estimate the joint development of both types of skill. It is not suitable to use separate models for each skill.

Finding 5: *Parental investment is multidimensional.* Investment must be included as a vector in the model. The components in this vector must distinguish between financial resources and other types of parental input. The choice of components must be driven by theories of child development and statistical analysis of the data.[41]

Finding 6: *Skills are self-productive.* Skill in one period must be defined as a function of the same type of skill in previous periods.

Finding 7: *Skills can be cross-productive.* Skill must also be defined as a function of the other types of skill in previous periods.[42]

I argue that adapting an existing human capital *production function*[43] allows me to define a model of skill formation which captures these key features of the literature. The remainder of this section explains the basic structure of this production function with specific focus on the modifications I have made to the existing model and justification for choosing this model specification. Understanding the motivation behind the model is important as it provides context for the details presented later in this chapter.

[41] The specifics of defining this vector are discussed in Section 3.3.

[42] Skills may not be cross-productive in all cases: if the skill in question is not cross-productive, this modelling strategy results in a co-efficient of zero on the lagged measure of the other skill.

[43] As discussed in the Literature Review: a *production function* defines a specific *technology* that gives the maximum output produced by a given vector of inputs. In the case of child development the inputs are prior skill, parental investment and environmental factors, while the outputs are current skill.

Modelling Skill Development Using a Production Function

This subsection introduces the specification of the human capital production function used in this thesis to model skill development in children. By defining the basic components of this production function I am able to show how this particular specification satisfies the seven key findings from the literature. Before I continue, it is important to note that this section only provides an introduction to the model and a discussion of why it is suitable for my research. Later in this chapter, in Section 3.2, I present the full details of how to mathematically specify this production function along with the specifics of how it will be used later in this thesis, in Chapter 4 and Chapter 5 to empirically estimate skill development.

The reader may recall from Section 2.3 that there are three main considerations when defining a human capital production function. These are: how to define human capital; what factors of production to include in the model; and which functional form the production function should take. In the context of using a human capital production function to mathematically model the relationship between parental inputs and childhood ability, these three considerations are recast as: how to define childhood ability[44]; how to define parental inputs to child development (e.g. parental behaviours and household resources); and how to mathematically model the relationships between these inputs and ability (model specification.) Each of these considerations is discussed below.

Defining Human Capital in Children

The first consideration when specifying the production function for skill development, is how to define human capital in children. This requires identifying the relevant childhood skills, and deciding how specify them within the model. The literature presented in Chapter 2 provides substantial evidence that a multitude of skills are linked with academic and life success. In order to capture the diverse nature of childhood abilities, this thesis follows the convention of labour economists who define human capital as a joint function of *cognitive* and *non-cognitive* skills.[45]

After identifying human capital as a joint function of these two skills, the next consideration is how to specify this function within the model. As discussed in Chapter 2, there is no consistent strategy for whether childhood ability should be defined as unidimensional, and be modelled using a scalar variable, or multidimensional, with a vector used to represent a set of abilities. Put differently, the model can measure a single type of ability (i.e. only considering cognitive ability) or it can define ability as multiple distinct constructs (i.e. considering both cognitive ability and non-cognitive ability).

[44] The reader is reminded that this thesis uses the terms ability and skill interchangeably. More information on this choice of terminology is presented in Section 2.1.

[45] The shape of this joint production function can take many functional forms. The necessary conditions for this functional form will be discussed later in this chapter.

As specified in Finding 4, the literature provides substantial evidence that *cognitive and non-cognitive skills cannot be examined in isolation.* Therefore, I follow the lead of Cunha and Heckman (2007) and present a model of development which defines childhood ability as a two-dimensional vector consisting of cognitive and non-cognitive ability. Compared to models which focus on a single type of ability, this modelling approach allows me to present more nuanced estimates of skill development in childhood.

Defining Parental Input to Child Development

When defining the production function for skill development, the second major consideration is how to specify the inputs to child development.[46] Chapter 2 has already highlighted the development of childhood skills is influenced by a multitude of factors. Since this thesis aims to understand the role of parental inputs, the present discussion focuses on the parental behaviours and traits which inform my specification of parental investment and does not discuss the other inputs to development that are included in the model as controls. While other factors of production are included in the model, their specification follows standard practice and will be addressed later in this chapter.

Within production functions in general, factors of production can be specified using a scalar variable or a vector. This is because, parental input can be either be conceptualised as unidimensional or as multidimensional. In other words, parental investment can be thought of as one general parenting factor or, instead, as multiple distinct 'types' of parenting inputs (i.e. separately considering financial resources, parental time, education related inputs, etc.). If an object is treated as multidimensional then it is possible to separately measure the role played by each distinct dimension. For example, if parental time is measured separately from financial resources then a model can capture the separate impact that each of these constructs has on development. Alternatively, if they are included as a combined construct of 'parental investment', then the model can only measure their joint effect and cannot identify the individual impact of each dimension.

In order to satisfy the literature which is captured by Finding 5: *"parental investment is multidimensional"*, I present an updated specification for the technology of skill formation that allows for multiple, distinct types of parental input.[47] More specifically, while Cunha and Heckman's original technology of skill formation assumes a single underlying parenting factor, and models parenting using a scalar variable, my updated model specifies that parenting is included as a vector. This allows me to differentiate between the effects of parental behaviour and attitudes as compared to family socio-economic resources. Defining parenting in this way allows me to examine a variety of research questions about the impact of parenting on child development.

[46] In the context of the production function, these inputs are known as *factors of production.*

[47] This is in line with the recommendations presented by Heckman and Mosso (2014) who suggest that future work should explore multidimensional models of parental investment.

Although human capital literature provides strong support for the definition of ability as a function of cognitive and non-cognitive skill, the specific constructs to include in the vector for parental investment are not as clearly defined by the literature. Therefore, the identification of parental inputs is a key feature of my theoretical framework and Section 3.3 is dedicated to explaining the methodology used to define this vector. For the purpose of defining the structure of the model, the specification of the vector is irrelevant, as long as the model is specified using an $n \times 1$ vector of investments, where $n \geq 1$. Further details of this specification are presented in the next section.

Choosing the Correct Specification for the Production Function

Once the dimensionality of the variables has been defined, the final consideration is how to specify the structural model in order to best capture the relationship between the child's skills and parental investment. This model can take many forms; however, as discussed in the previous chapter, human capital production functions provide the most mathematically developed model of skill formation. While the mathematical structure of production function models is based on an economic theoretical framework, I propose a specification of the production function which uses measures taken from the psychometric literature and defines the relevant inputs to the production function based on theoretical constructs of parenting and child development from the field of developmental psychology.

The literature review has already shown that *value-added with lagged inputs technology of skill formation*[48] is the best way to define a production function for the development of cognitive ability. By definition, this type of production function satisfies Finding 1, Finding 2, Finding 3 and Finding 6 from the literature review. Below, I discuss how using the definitions of human capital and parental investment that I have already introduced allow me to modify this production function to satisfy the remaining three key findings.

To begin, I focus on Finding 4: *cognitive and non-cognitive skills cannot be examined in isolation*. As described above, defining human capital as a two-dimensional vector of childhood ability is the best way to address this finding. Specifically, I adopt the model specification introduced by Cunha and Heckman (2007). The authors present a *value-added with lagged inputs technology of skill formation* where childhood ability is defined as a two-dimensional vector. This model specification allows for an estimation of the joint evolution of cognitive and non-cognitive skills. Furthermore, specifying the model in this way means that the vector of skills in one period is a function of the vector of skills in the previous period, thereby satisfying Finding 7: *skills can be cross-productive*.

Finally, while the specification presented by Cunha and Heckman (2007) is well suited to capturing the connected development of cognitive and non-cognitive skills, it

[48]This is a recursive model wherein skills are formed over multiple periods and the skills 'outputted' from one period — in addition to other factors from the child's environment — are the inputs for the next period.

needs to be modified in order to explain the multiple aspects of parenting described by Finding 5: *parental investment is multidimensional.*[49] To achieve this, I propose using the definition of parental investment provided above, and adapting Cunha and Heckman's (2007) production function to include parental investment as a vector. This results in a *value-added with lagged inputs, multidimensional specification of ability and investment, technology of skill formation* which simultaneously estimates the trajectory of cognitive and non-cognitive ability using multiple types of parental input as factors of production.

This model of skill formation addresses all the key conclusions from the literature on child development and parenting that were summarised in Section 2.6 and it allows me to answer the research questions introduced in Chapter 1. Full details of how to estimate this model are included in the next section.

[49] That is not to say that these researchers were not aware of this aspect of skill development. In their overview of recent literature on skill development, Heckman and Mosso (2014) discuss the need for models to estimate multiple types of parental input but only reference works that examine multiple types of parenting using different methodological approaches.

3.2 Full Model Specification

Empirically estimating this modified technology of skill formation requires the use of a dynamic model consisting of two components: a structural model and a set of measurement models. The structural model is the mathematical representation of the technology of skill formation, whereby a child's skills evolve according to a law of motion influenced by parental inputs and environmental factors. The structural model cannot be estimated directly because parental investment and child skill are unobservable latent constructs. Instead, using an approach proposed by Cunha and Heckman (2008), a set of measurement models is defined to estimate these underlying latent constructs from observable indicators. The structural model and measurement models are combined in a structural equation model which estimates all of the parameters simultaneously.

Each component of the model is discussed separately below, along with the conditions necessary for identification of the full model.[50]

Structural Model

To begin, let the vector $\theta_t = \left(\theta_t^C, \theta_t^N\right)'$ represent the stock of latent cognitive skills θ^C and latent non-cognitive skills θ^N of the child at time t. A child's ability in the next period is given by θ_{t+1} and is a function of: their prior skill, θ_t; parental investment in the prior period, I_t; and observable exogenous measures of socioeconomic status, X_t. Thus, the evolution of skill over time can be expressed using the recursive function

$$\theta_{t+1} = f\left(\theta_t, I_t, X_t\right), \tag{3.1}$$

where f is a production function, as introduced Section 2.3. Similar to the production functions for physical goods, the technology of skill formation can take many forms.

Imposing several constraints on f guarantees that the model captures key elements of the empirical evidence on skill development. First, f is assumed to be increasing in θ_t, I_t and X_t. This condition allows for positive marginal returns to all inputs, which ensures that prior ability, parental input and socio-economic factors all promote the development of future ability. Next, the function is assumed to be concave in I_t. This concavity implies that the marginal returns of skill to parental input are either constant or decreasing. Thus, as the level of investment increases, the rate of skill development does not rise. Finally, f is assumed to be twice continuously differentiable in all arguments, which

[50] As the modified technology of skill formation and the strategy of using measurement models both draw heavily on the works of Cunha and Heckman (2007, 2008), there is substantial overlap with their original works in my exposition of the model used in this thesis. In the text that follows, I do not cite each equation directly, but it is to be assumed that the mathematical models are based on their original work, with the exception of the specified modifications.

allows for both the first and second derivative to be identified at all points. Under some circumstances the assumption of twice differentiability can be relaxed, but it is retained for mathematical convenience.

There are numerous functions which meet the above requirements, but Cunha and Heckman (2008) and Cunha et al. (2010) have provided evidence that using a linear function for f provides a good basis for measuring the stage-specific effects of parental investment. These stage-specific estimates allow for the identification of critical and sensitive periods of investment. Unfortunately, since a linear model is separable in its inputs, the substitutability of inputs cannot be calculated. As a result of this limitation, Cunha et al. (2010) proposed the use of non-linear models for measuring the inter-temporal substitutability of investment, and the static complementarity of skill and investment. Fortunately, the primary objective of my theoretical model is to examine the stage-specific impacts of different types of parenting; this can be accomplished using a linear model, therefore the complexity of the non-linear models is not required.

Since a linear specification is far less computationally and data-intensive, and the measures of substitutability and complementarity are not of specific interest, a linear model is more than sufficient for the present framework. Using this approach avoids some of the statistical considerations and large sample size requirements that must be taken into account when using non-linear models.[51]

Assuming that f is linear, evolution of a child's ability over time is given by

$$\theta_{t+1} = \Gamma_t \theta_t + B_t I_t + \Lambda_t X_t + \eta_t, \tag{3.2}$$

for $t \in 1, \ldots, T$, where θ_t is the (2×1) vector of skills defined above, I_t is a $(s \times 1)$ latent vector of parental investments, X_t is an observed matrix of exogenous variables, Γ_t, B_t and Λ_t are time varying parameters to be estimated by the model and η_t is the error term, assumed to be independent across individuals and over time. The process in Equation 3.2 can be rewritten by repeatedly substituting in θ_t therefore expressing θ_{t+1} as a function of initial ability along with the entire history of parental inputs. This is consistent with skill development being a cumulative process.

Linear Laws of Motion for Each Type of Skill:

Equation 3.2 can also be represented using a system of equation with two separate linear laws of motion for non-cognitive and cognitive skills. These can be written as:

$$\theta_{t+1}^N = \gamma_{1,t}^N \theta_t^C + \gamma_{2,t}^N \theta_t^N + \beta_{1t}^N I_t^1 + \cdots + \beta_{st}^N I_t^s + \Lambda_t^N X_t + \eta_t^N, \tag{3.3}$$

[51] For further detail on the differences between linear and non-linear specifications, as well as the mathematical proofs of each specification of the model, the reader is directed to the discussion of both models provided by Heckman and Mosso (2014).

and

$$\theta_{t+1}^C = \gamma_{1,t}^C \theta_t^C + \gamma_{2,t}^C \theta_t^N + \beta_{1t}^C I_t^1 + \cdots + \beta_{st}^C I_t^s + \Lambda_t^C X_t + \eta_t^C, \tag{3.4}$$

respectively. Note that although the variables included in both laws of motion are the same, the estimated parameters will be different for each type of skill.

Although the laws of motions can be represented separately, these parameters must estimated simultaneously.Simultaneous estimation calculates the parameters that are the best solution for the joint system of equations, whereas separately estimating the parameters for each type of skill treats the other type of skill as fixed and therefore is only finding the solution to a single equation. As the two types of skills are interdependent, varying the parameters in one equation will change the nature of the second equation. Estimating the skills separately will completely neglect this feature of the model. For example, if maximum likelihood estimation is used, a simultaneous estimation will estimate a set of parameters that maximises the likelihood of seeing the recursive relationship across both types of skills, assuming that both cognitive and non-cognitive skills are endogenous in the model. If the equations are estimated separately, then a maximum likelihood estimation only maximises the likelihood of seeing the recursive relationship in one type of skill, treating the other as fixed. Because simultaneous and separate estimation models are not measuring the same thing, they can produce vastly different estimates for the same parameters.

Specifying the Child's Starting Position:

As with any recursive model, this production function requires assumptions regarding the model's first period. It is assumed that the child is born with an initial endowment of skill θ_0, determined by environmental and genetic factors X_0^θ, where θ_0 can be expressed as

$$\theta_0 = \psi_0 X_0^\theta + \xi_0, \tag{3.5}$$

where ψ_0 is a matrix of estimated parameters and X_0^θ contains the period-specific measures included in X_t as well as time-invariant demographic characteristics used to capture family background[52], health at birth[53] and postnatal factors which indicate early child health. Though the theoretical model assumes the first period of the model corresponds with the first period of development, the model can be modified to start at any point in a child's development if early data is not available. Changing the starting period does not change any of the model assumptions but does fail to differentiate between skill gained from early childhood investments and the child's endowment of skill at birth.

[52]These include measures such as parental ethnicity, the mother's age at the time of the child's birth, and the native language(s) of the child's parents.

[53]This is commonly captured by the child's birth weight.

Modifying the Model to Include Multiple Investment Types:

While Cunha and Heckman (2007) defined parental investment as a single latent factor, the present analysis modifies the production function to define I_t as a vector (i.e. $s \neq 1$). This approach is complementary to recent work by Attanasio et al. (2015) which defined a multidimensional set of parental inputs. However, while Attanasio et al. (2015) specified separate production functions for cognitive skills and socio-emotional skills, I specify them using a joint production function. More precisely, as discussed in Section 3.1, I define human capital using a vector of cognitive and non-cognitive ability; this vector evolves according to a single production function.

By defining I_t as a vector, the model explicitly allows for the possibility of multiple factors driving parental investment, each of which can have unique impacts on skill development. The s types of parental investment are representative of different types of parenting behaviours and each type of investment is measured by different underlying factors and has a separate effect on skill development. The separation of parental investment into different factors allows for the identification of the specific impact of different types of parenting skills. Section 3.3 goes into detail about how these multiple types of parental investment are defined, but for the purpose of defining the model it is enough to know that there are multiple, distinct types of parental investment.

The Effect of Covariates for Investment:

Though the law of motion for skill presented in Equation 3.2 captures the direct effect of covariates on skill formation, the structural model must also define the impact that exogenous covariates have on parental investment. It is assumed that, although there are separate types of parental investment, all these types are influenced by certain common factors. These factors are represented by a matrix of observable variables X_t^I. The matrix X_t^I is related to the vector of parental investments by the function

$$I_t = \phi_t^I X_t^I + \varsigma_t, \tag{3.6}$$

where ϕ_t^I is a matrix of estimated parameters and ς_t is the error term, independent across individuals and over time. Though there is some overlap, the matrix of observed variables X_t^I, does not necessarily contain the same observed variables as X_t. Factors that directly affect parental investment are included in X_t^I, while X_t includes factors which directly impact ability. For example, parental investment, I_t, is influenced by the number of siblings because parents in larger families must split their time among children; as such, family size is included in X_t^I in Equation 3.6. However, family size is not included in X_t in Equation 3.2 as the number of siblings does not directly impact a child's ability.

Representing the Structural Model Using a Path Diagram:

Although linear laws of motion provide a mathematical representation of the structural model, it is helpful to illustrate the recursive model using a path diagram. A visual representation allows the reader to see the extent to which the laws of motion are interconnected for the two types of skills.

Figure 3.1 depicts the first four periods of the structural model with each time period represented by a different colour. Following the standard notation for structural equation models, observable variables are represented using rectangles and latent variables represented using ovals.

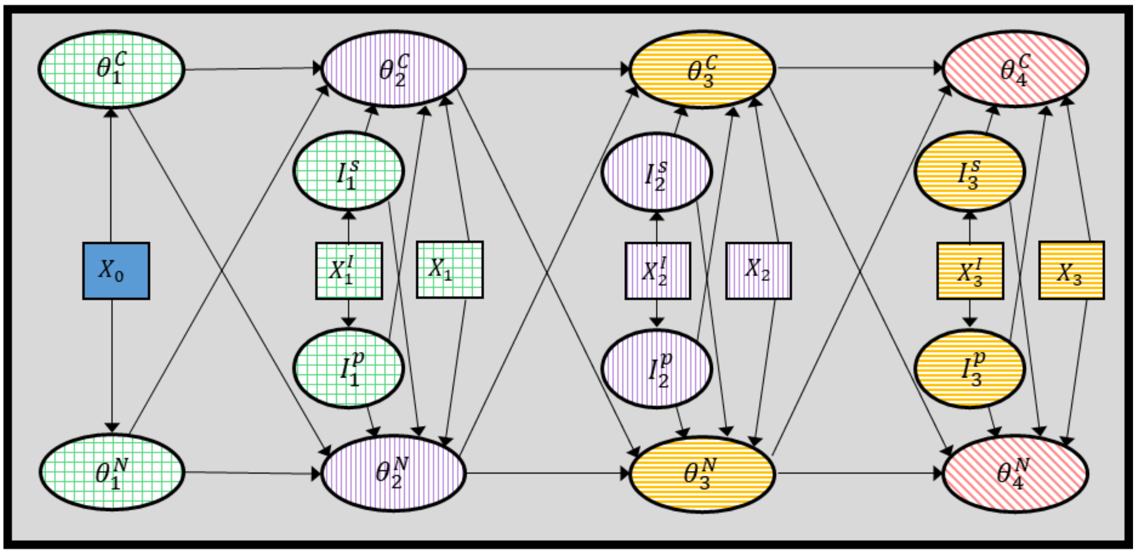


Fig. 3.1 Structural Model of Skill Formation with Two Investment Types

This diagram demonstrates the complex set of relationships that are included in the two laws of motion described earlier in this section. In particular, the path diagram illustrates how skills in one period are the product of skills and inputs in all prior periods. Unless repeated substitution is used to expand a given period of the structural model, this recursive nature is less readily apparent from the laws of motion.

The visual representation provided by Figure 3.1 also emphasises how interconnected the two laws of motion become once development of these two skills is modelled over multiple periods. Since the model simultaneously estimates parameters for all the arrows in this diagram, it is easy to see how modelling the production function separately for each type of skill might yield very different results from the simultaneous specification.

Implications of this Functional Form:

The structural model defined above identifies stage-specific parameters for the effect of each type of parental input on the development of cognitive and non-cognitive skills. The model also estimates the lagged effect of prior skill and the period-specific impacts of relevant covariates. Not only does each parameter precisely measure one aspect of development, but taken together they offer a clearer understanding of child development.

Furthermore, this structural model also allows researchers to mathematically define several other concepts that are often used when discussing childhood skill formation. Each of these concepts is based in existing empirical work, but by taking the first derivative of Equation 3.2, it is possible to measure the self-productivity and cross-productivity of each type of skill, and to identify critical and sensitive periods for parental investment. The specifics of these measures are discussed below.

Self-Productivity: The model requires that skill be self-productive, but the level of self-productivity varies by period. By definition, self-productivity implies that prior skill has a positive marginal product. Mathematically, this is given by

$$\Gamma_t^k = \frac{\partial \theta_{t+1}^k}{\partial \theta_t^k} > 0, \tag{3.7}$$

where the magnitude of self-productivity is equal to the coefficient Γ_t^k in Equation 3.2.

Cross-Productivity: Similarly, the model is defined to allow for skills to be cross-productive. This implies that one type of skill fosters the development of the other type of skill. If cross-productivity occurs, it will be given mathematically by

$$\Gamma_t^l = \frac{\partial \theta_{t+1}^l}{\partial \theta_t^k} > 0 \text{ for } l \neq k, \tag{3.8}$$

where the magnitude of cross-productivity is equal to the coefficient Γ_t^l in Equation 3.2.

Sensitive Period: Unlike cross-productivity and self-productivity, which only examine the parameters from a single period, sensitive periods of investment require measurements from all other periods of the model. A sensitive period of investment occurs when investment is more productive in the given period compared to all other periods. Since sensitive periods are the most efficient time to invest, identifying such periods is integral for creating targeted policy recommendations. Mathematically, a sensitive period of investment occurs when

$$\frac{\partial \theta_{t+1}}{\partial I_{t^*}} > \frac{\partial \theta_{t+1}}{\partial I_s} \text{ for periods } t^* \neq s, \tag{3.9}$$

with each type of investment having one sensitive period of investment for a given skill.[54]

[54] It is possible to have multiple sensitive periods, if the marginal product of investment is equal across several periods.

Critical Period: As discussed in Chapter 2, there is ongoing debate about the existence of critical periods, as they imply that investment is only effective during a single period of development and has no impact on skill formation outside of this period. Within an economic production function, a critical period of investment for a given skill occurs when the marginal product of investment is positive for the specified period but is zero in all other periods.[55] Mathematically, this is represented as

$$\frac{\partial \theta_{t^*+1}}{\partial I_{t^*}} > 0 \text{ for } t^* \text{ but } \frac{\partial \theta_{t+1}}{\partial I_t} = 0 \text{ for } t \neq t^*. \tag{3.10}$$

Defining the model in this way allows for identification of critical periods if they do exist, but also implies that they do not need to occur in a given empirical application. This flexibility is important, because although there is some evidence to support the existence of critical periods for specific aspects of development, there is substantial evidence that for the majority of skills, there are multiple periods with positive returns to investment.[56]

[55] By this definition, a critical period of investment will have a larger marginal product of investment than any other period, and thus will also be a sensitive period of investment.

[56] See Howard-Jones, Washbrook, and Meadows (2012) for a discussion of why sensitive periods are more likely than critical periods.

Measurement Model

As discussed in the literature review, it is assumed that the child's cognitive skills, their non-cognitive skills, and the multiple types of parental investment are all latent variables and cannot be directly observed. Using dynamic factor models, as proposed by Cunha and Heckman (2008), these latent variables can be represented within the structural model using measurement models that are a function of observable indicators. Simultaneous estimation can then be used to estimate the latent variables and the parameters of interest in this combined system of equations.

Before estimating the full dynamic factor model, the individual measurement models for each latent factor must be defined. To begin, let θ_t^C, θ_t^N and $\theta_t^{I_1} \ldots \theta_t^{I_s}; s \in R$ represent the latent constructs of cognitive skills, non-cognitive skills and parental investment types.[57] Though it is not possible to directly observe the latent variables (θ_t^k), for each $k \in \{C, N, I_1, \ldots, I_s\}$, there are multiple observable indicators ($Y_{j,t}^k; j \in \{1, \ldots, m_t^k\}$), which each contain some information about the corresponding latent variable. The observable indicators are themselves a function of the latent variable. The total number of indicator variables for skill θ_t^k is given by m_t^k, with $m_t^k \geq 2$. In the existing literature, one commonly used approach is to combine measures using indices in which all measures are given an equal weight. Though an ad-hoc combination of these measures could be used to create indices[58], such an approach introduces potential for substantial bias as a result of constructing the index using poorly suited weights.[59]

Instead of imposing a structure to calculate an ad-hoc index for the latent variables, a system of equations is created to estimate the index that best captures the differing amount of information captured by each indicator. To accomplish this, a measurement model is created where each of the observable indicators of θ_t^k is modelled as

$$Y_{j,t}^k = \mu_{j,t}^k + \alpha_{j,t}^k \theta_t^k + \Phi_{j,t}^k Z_t^k + \varepsilon_{j,t}^k, \tag{3.11}$$

where there are m_t^k equations, one for each of the indicator variables. Each indicator is an imperfect measure of the underlying latent factor and is influenced by other observable and unobservable variables. To model some of this measurement error, each equation includes Z_t^k, a matrix of covariates known to influence the measured indicator but independent of the underlying latent factor. Assuming that the measurement errors are not correlated across the measures, but rather are normally distributed and equal to the standard deviation of the observed indicators $\left(\varepsilon_{j,t}^k = \rho_{k,j,t}^2\right)$, the correlation between the indicators ($Y_{j,t}^k$) can be attributed to the underlying latent variables (θ_t^k) and the observed covariates (Z_t^k). By normalising $\alpha_{1,t}^k = 1$, it is possible to estimate the remaining

[57] For notational ease, the remainder of this section rewrites $I_t^1 \ldots I_t^s, s \in R$ as $\theta_t^{I_s}, s \in R$.

[58] The most commonly used technique is to calculate an average of a set of test scores.

[59] See Cunha and Heckman (2008) for a review of the limitations of ad-hoc indices.

parameters $\alpha^k_{j,t}$; these are known as factor loadings.

For example, true cognitive ability cannot be directly observed, but the separate scores from an array of cognitive tests each provide some information about the underlying level of cognitive skill. A child's score on a given cognitive test is assumed to be an imperfect measure of their underlying cognitive ability: the test is a biased measure as it contains measurement error from test design, alongside measurement error capturing systemic biases in how the test measures cognitive ability in certain demographic groups. While it is possible to use the observed covariates Z^C_t to control for the observable characteristics that are known to correlate with cognitive test scores, other measurement error remains. With only one test score, it is impossible to know how much of the test score is attributable to true ability, θ^C_t, and how much is simply bias. If instead there are scores for three separate cognitive tests taken during the same period — and we assume that the remaining measurement error of the three tests is uncorrelated — it is possible to use these three measures to identify true ability θ^C_t. By constructing Equation 3.11 for each of the three tests, the following system of equations is created:

$$\begin{aligned} Y^C_{1,t} &= \mu^C_{1,t} + \alpha^C_{1,t}\theta^C_t + \Phi^C_{1,t}Z^C_t + \varepsilon^C_{1,t} \\ Y^C_{2,t} &= \mu^C_{2,t} + \alpha^C_{2,t}\theta^C_t + \Phi^C_{2,t}Z^C_t + \varepsilon^C_{2,t} \\ Y^C_{3,t} &= \mu^C_{3,t} + \alpha^C_{3,t}\theta^C_t + \Phi^C_{3,t}Z^C_t + \varepsilon^C_{3,t}. \end{aligned} \tag{3.12}$$

This system of equations has three equations which each model the observed indicator $Y^C_{j,t}$ as a function of the observable covariates Z^C_t, and the unobservable θ^C_t. By setting $\alpha^C_{1,t} = 1$ and estimating these three equations, it is possible to isolate and solve for θ^C_t.

Special Case for Ordinal Indicator Variables:

In some cases, the observable indicators used for the measurement model are reported using an *ordinal scale*, with responses given as integers representing one of the possible R^q_t response categories.[60] For example, a variable measured on a Likert scale is reported using five categories (i.e. $R^q_t = 5$): 'Strongly disagree', 'Disagree', 'Neither agree nor disagree', 'Agree' and 'Strongly agree'. For ordinal scales containing fewer than seven categories, treating the ordinal data as if it were continuous introduces considerable bias (Dolan, 1994; Lubke & Muthén, 2004; Rhemtulla, Brosseau-Liard, & Savalei, 2012).[61]

To avoid introducing this bias, ordinal variables with fewer than seven categories require a slightly modified version of the continuous measurement model described above; this is known as an *ordered logit model*. Although some readers will be familiar with ordered logit models, the specifics are briefly reviewed below to ensure clarity.

[60] An *ordinal scale* is a type of non-continuous variable reported using discrete categories (Fraenkel & Wallen, 2009). This is also referred to as ordered categorical data.

[61] When categorical variables have "seven or more categories the bias [from treating the variable as continuous] is very small"(Hox, Moerbeek, & van de Schoot, 2017, p.130).

In this thesis, the latent constructs that are to be estimated using ordinal indicators are denoted as: θ_t^q, for each $q \in \{C, N, I_1, \ldots, I_s\}$. For each θ_t^q, there are m_t^q observable ordinal indicators, with $m_t^q \geq 2$. These indicators are given by $Y_{p,t}^q$, $p \in \{1, \ldots, m_t^q\}$, with each observation of $Y_{p,t}^q$ taking on a value $r = 1, \ldots, R_t^p$, $r \in \mathbb{Z}$.

In a continuous measurement model, as described above, it is assumed that the observed indicators $(Y_{p,t}^k)$ can take on any real number; each indicator variable is therefore directly modelled as $Y_{j,t}^k = f(\theta_t^k, Z_t^k)$, which is a linear function of the unobservable latent factors. Since ordinal variables can only take on integer values, an adjustment has to be made so that the measurement equations are specified so that $Y_{p,t}^q$ is defined as an ordinal variable obtained from some function of the unobservable latent factors. To do this, ordered logit models assume that a continuous unreported variable $(Y_{p,t}^q)^* = f(\theta_t^q, Z_t^q)$ underlies the reported ordinal response variable $Y_{p,t}^q$. Thresholds along the continuous distribution of $(Y_{p,t}^q)^*$ determine which value of the ordinal variable $Y_{p,t}^q$ is reported.[62] Using this logical framework, linear equations can be created to model the continuous measures $(Y_{p,t}^q)^*$ underlying each of the reported ordinal measures $Y_{p,t}^q$.

In the context of the present model, the ordinal regression estimates a function for $(Y_{p,t}^q)^*$, which is used to define the cut-points, ρ_r^p, for each of the categories of the ordinal score $Y_{p,t}^q$. For example, for a predicted $(Y_{p,t}^q)^*$ a reported Likert score would be given by:

$$Y_{p,t}^q = \begin{cases} 1 \Rightarrow \textit{Strongly disagree} & \text{if } \rho_0^p \leq (Y_{p,t}^q)^* < \rho_1^p \\ 2 \Rightarrow \textit{Disagree} & \text{if } \rho_1^p \leq (Y_{p,t}^q)^* < \rho_2^p \\ 3 \Rightarrow \textit{Neither agree nor disagree} & \text{if } \rho_2^p \leq (Y_{p,t}^q)^* < \rho_3^p \\ 4 \Rightarrow \textit{Agree} & \text{if } \rho_3^p \leq (Y_{p,t}^q)^* < \rho_4^p \\ 5 \Rightarrow \textit{Strongly agree} & \text{if } \rho_4^p \leq (Y_{p,t}^q)^* < \rho_5^p \end{cases} . \tag{3.13}$$

Therefore, the ordinal measurement mode for the latent constructs takes the form:

$$(Y_{p,t}^q)^* = \mu_{p,t}^q + \alpha_{p,t}^q \theta_t^q + \Phi_{p,t}^q Z_t^q + \varepsilon_{p,t}^q \tag{3.14}$$

such that $Y_{p,t}^q = r$ if $\rho_{r-1}^p \leq (Y_{p,t}^q)^* < \rho_r^p$ where $\rho_0^p = -\infty$ and $\rho_{R_t^q}^p = \infty$. As with the continuous case, each measurement equation includes the matrix of covariates Z_t^q that are known to influence the measured indicator but are independent of the underlying latent factor. Assuming that the measurement errors are normally distributed and the standard deviation of the observed indicators is given by $(\varepsilon_{p,t}^q = \sigma_{q,p,t}^2)$, then the correlation between the indicators $(Y_{p,t}^q)$ can be attributed to the underlying latent variables and the observed covariates Z_t^q. As with the continuous measurement models, it is only by normalising $\alpha_{1,t}^q = 1$ that it is possible to estimate the remaining factor loadings $\alpha_{p,t}^q$.

The model fits values of ρ_r^p so that when $(Y_{p,t}^q)^*$ crosses the cut-point, the observed category $Y_{p,t}^q$ changes. The statistical model attempts to maximise the likelihood of the

[62] Hedeker (2008) provides more information about this approach to modelling categorical variables.

predicted cut-points ρ^p_r, yielding the reported ordinal scores. A full statistical explanation of this type of model is beyond the scope of this thesis. For further details the reader is directed to Anderson, Kim, and Keller (2014) Hayashi (2000) or Hedeker (2008).

Representing Measurement Models using Path Diagrams:

Each of the systems of equations described above can be represented visually using a path diagram. An example measurement model for cognitive ability is illustrated in Figure 3.2. Similar diagrams can be created for each latent construct in the model.

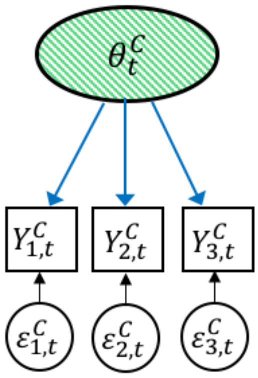


Fig. 3.2 Example of a Measurement Model for Cognitive Ability

Using Factor Loadings to Estimate Latent Factor Scores

The estimated factor loadings $\hat{\alpha}^k_{j,t}$ and $\hat{\alpha}^q_{p,t}$ can be used to calculate estimated scores for the latent variables. For the present framework, these estimated latent scores are purely for reference, as the full model is estimated simultaneously and these calculations take place within the model.

The process for calculating the latent factor scores is as follows: first, the Thurstone (1959) procedure defines implicit relative weights $w_{j,t}$ as

$$w_{j,t} = \frac{\left(\hat{\alpha}^k_{j,t}\right)^2}{\sum\limits_{j=1}^{m^k_t}\left(\hat{\alpha}^k_{j,t}\right)^2}. \tag{3.15}$$

Using these implicit weights, a weighted average of the observable indicators can be obtained. This weighted average is the error-corrected estimate of the underlying latent ability, which is also known as the latent factor score and can be expressed as

$$\hat{\theta}^k_t = \sum_{j=1}^{m^k_t} w_{j,t} Y^k_{j,t}. \tag{3.16}$$

Full Set of Measurement Models

For each of the unobservable latent variables included in the structural model, a separate measurement model is required. These distinct sets of measurement equations combine to give the full set of measurement models that inform the structural model described above.

The final specification for the full set of measurement equations will vary depending on how many latent constructs require measurement models and on whether the indicators for a given construct are ordinal or continuous. If, for example, all of the indicators are continuous, then the full set of measurement models is given by:

$$\begin{aligned} Y_{j,t}^{C} &= \mu_{j,t}^{C} + \alpha_{j,t}^{C}\theta_t^{C} + \Phi_{j,t}^{C} Z_t^{C} + \varepsilon_{j,t}^{C} \\ Y_{j,t}^{N} &= \mu_{j,t}^{N} + \alpha_{j,t}^{N}\theta_t^{N} + \Phi_{j,t}^{N} Z_t^{N} + \varepsilon_{j,t}^{N} \\ Y_{j,t}^{I_1} &= \mu_{j,t}^{I_1} + \alpha_{j,t}^{I_1}\theta_t^{I_1} + \Phi_{j,t}^{I_1} Z_t^{I_1} + \varepsilon_{j,t}^{I_1} \\ \vdots & \qquad\qquad \vdots \\ Y_{j,t}^{I_s} &= \mu_{j,t}^{I_s} + \alpha_{j,t}^{I_s}\theta_t^{I_s} + \Phi_{j,t}^{I_s} Z_t^{I_s} + \varepsilon_{j,t}^{I_s}, \end{aligned} \tag{3.17}$$

with each measurement equation being a linear model. If however, all indicators are ordinal, then the full set of measurement models is instead:

$$\begin{aligned} (Y_{p,t}^{C})^* &= \mu_{p,t}^{C} + \alpha_{p,t}^{C}\theta_t^{C} + \Phi_{p,t}^{C} Z_t^{C} + \varepsilon_{p,t}^{C} \\ (Y_{p,t}^{N})^* &= \mu_{p,t}^{N} + \alpha_{p,t}^{N}\theta_t^{N} + \Phi_{p,t}^{N} Z_t^{N} + \varepsilon_{p,t}^{N} \\ (Y_{p,t}^{I_1})^* &= \mu_{p,t}^{I_1} + \alpha_{p,t}^{I_1}\theta_t^{I_1} + \Phi_{p,t}^{I_1} Z_t^{I_1} + \varepsilon_{p,t}^{I_1} \\ \vdots & \qquad\qquad \vdots \\ (Y_{p,t}^{I_s})^* &= \mu_{p,t}^{I_s} + \alpha_{p,t}^{I_s}\theta_t^{I_s} + \Phi_{p,t}^{I_s} Z_t^{I_s} + \varepsilon_{p,t}^{I_s}, \end{aligned} \tag{3.18}$$

with ordered logit models being used. However, in most empirical applications, the observable indicators will be continuous for some latent constructs and ordinal for others. Therefore, the complete set of measurement models will be a composite of the two sets listed above.

For the remainder of this chapter, the measurement equations are used to describe a general concept under the assumption that the correct type of model will be used depending on whether the observable indicators are continuous or ordinal. Though the specific measurement equations differ, the use of measurement models within the structural model does not change depending on the type of indicator. For consistency in the text, the reader is to assume that $\varepsilon_{j,t}^{k}$ is synonymous with $\varepsilon_{p,t}^{q}$.

Combining the Structural and Measurement Models

The full structural equation model combines the equations for the structural model given by Equation 3.2, Equation 3.5, and Equation 3.6 with the equations for the relevant measurement models given by Equation 3.11 and Equation 3.14. This forms a system of equations that allows for the identification of the model and simultaneous estimation of all parameters.[63] Figure 3.3 illustrates two periods of the full model.[64]

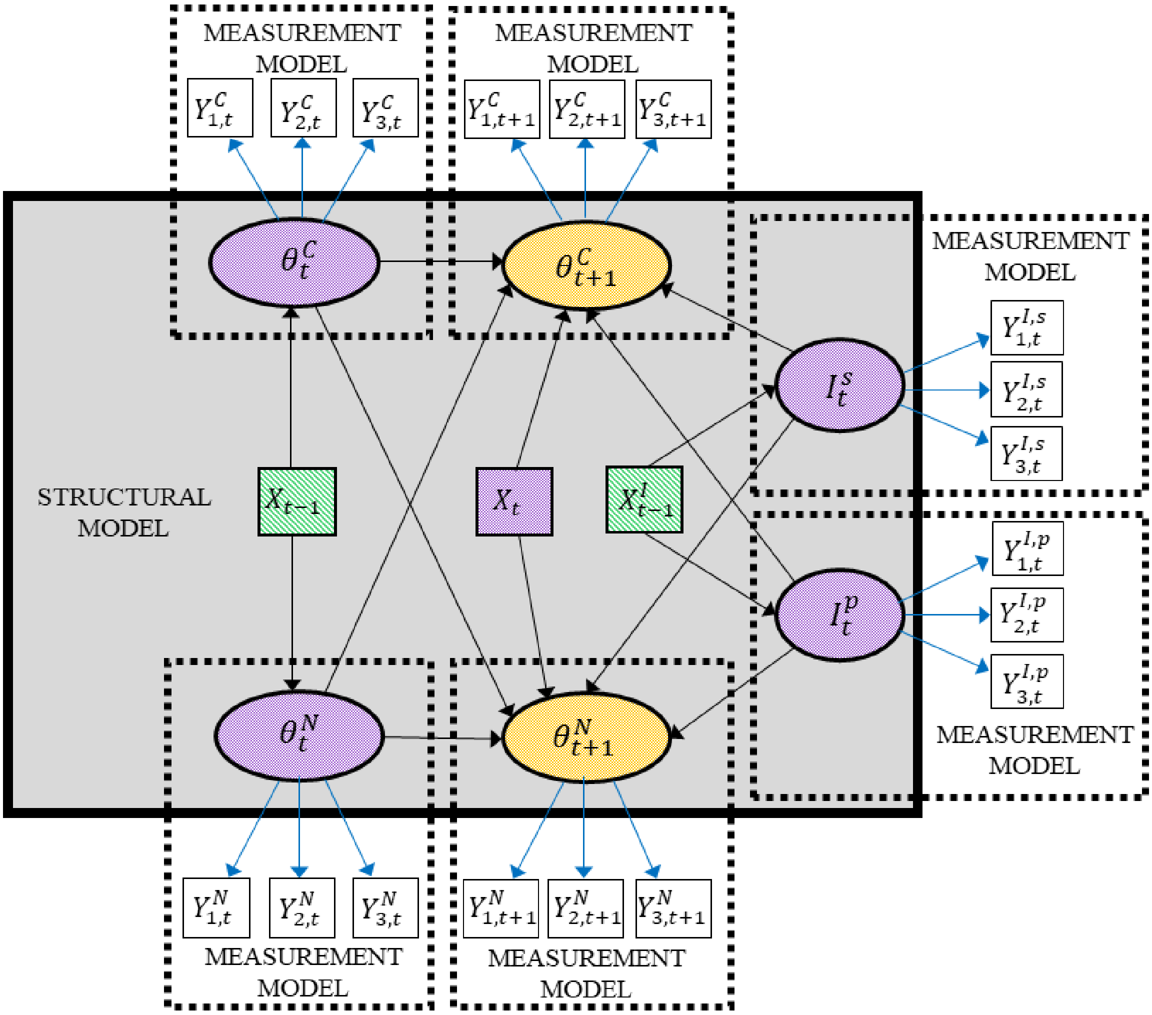


Fig. 3.3 Full Dynamic Model of Skill Formation with Two Investment Types

This two-period representation may be a simplified version of the multi-period model that would be estimated, but it nevertheless demonstrates how the two components combine to simultaneously estimate the model. When estimating the full model of child development, this path diagram would be extended to cover multiple periods.

[63] As discussed in Section 2.2, some researchers follow a two-step approach for estimating structural equation models. When using *factor score regression*, estimated factor scores are created in the first step and then substituted into linear regression models in the second step.

[64] The model in Figure 3.3 is specified with only two types of parental investment, but the diagram can be adjusted to include more types of investment.

Conditions Necessary for Identification of Full Model

The system of equations given by the structural model[65] and the measurement models[66] must be estimated simultaneously to identify the full model.[67] In order to identify all the relevant parameters, four assumptions must be made about the nature of the error terms in the equations for the measurement models, as well as one assumption the error terms in the equations for the structural model. These five assumptions are:

Assumption 1: $\varepsilon^k_{j,t}$ is mean zero and independent across agents and over time for $t \in \{0, \dots, T\}$; $k \in \{C, N, I_1, \dots, I_s\}$; $j \in \{1, \dots, m^k_t\}$; and $s \in \{1, \dots, S\}$.

Assumption 2: $\varepsilon^k_{j,t}$ is independent of $(\theta^C_\tau, \theta^N_\tau, I^1_\tau \dots I^s_\tau)$ for all $t, \tau \in \{0, \dots, T\}$; $j \in \{1, \dots, m^k_t\}$; $k \in \{C, N, I_1, \dots, I_s\}$; and $s \in \{1, \dots, S\}$.

Assumption 3: $\varepsilon^k_{j,t}$ is independent of $\varepsilon^l_{i,t}$ for $t \in \{0, \dots, T\}$; $l, k \in \{C, N, I_1, \dots, I_s\}$; $i, j \in \{1, \dots, m^k_t\}$ and $i \neq j$ if $k = l$.

Assumption 4: $\varepsilon^k_{j,t}$ is independent of $\varepsilon^l_{i,t}$ for $t \in \{0, \dots, T\}$; $i, j \in \{1, \dots, m^k_t\}$; $l, k \in \{C, N, I_1, \dots, I_s\}$ and $l \neq k$.

Assumption 5: η_t, ς_t and ξ_0 are independent of each other and independent of θ^C_t, θ^N_t, $I^1_t \dots I^s_t$, X_t, X^I_t, and X^θ_0 for all $t \in \{0, \dots, T\}$.

Assumption 1 states that for each measurement equation there is no relationship between the error term of one individual and the error term of another individual, or between the error term of an individual in one period and their error term in another period. Assumption 2 is the standard endogeneity assumption that the explanatory variables in the measurement equations are uncorrelated with the error terms. More specifically, Assumption 2 states that after controlling for the variables provided in each equation, the error terms in the measurement equations are not correlated with any of the other latent constructs included in the model. An example of this would be that an item used as an indicator of non-cognitive ability cannot have measurement error correlated with the child's cognitive ability. Assumption 3 states that the error terms of each measurement equations for a given unobserved latent variable are uncorrelated. For example, if two items used to measure parental input both describe socially undesirable parenting behaviours (e.g. physical discipline) it is likely that both items will be under-reported. This measurement error results from method used to assess the constructs and is correlated across the two tests. The measurement model will falsely attribute

[65] Equation 3.2, Equation 3.5, and Equation 3.6

[66] Equation 3.11 and Equation 3.14

[67] Though it is possible to first predict factor scores from the measurement models then substitute these into the structural model, estimation using a two-step approach is subject to significant bias. For further detail refer to Skrondal and Laake (2001).

this correlated measurement error as shared variance relating to the latent construct. Assumption 4 states that the error terms from one measurement equation do not correlate with those for another measurement equation for a different unobserved latent variable. This means that the correlation among the measurement variables can only be attributed to the common effects of the unobserved latent variables (θ_t^k, I_t^s) and the observable covariates, and that any unexplained components of the measurement variables are independent across measurement models. For example, according to Assumption 4, if measures of cognitive ability and non-cognitive ability are both administered in English, they will likely underestimate the ability of a child who does not speak any English. If the model includes language ability as an observed covariate in each measurement equation then the measurement error of the two sets of tests will not be correlated. On the other hand but if language ability is not controlled for, then Assumption 4 will be violated as $\varepsilon_{j,t}^k$ is no longer independent of $\varepsilon_{i,t}^l$. Finally, Assumption 5 is the standard classical measurement error assumption that within the structural model errors are independent of all other variables.

These assumptions are an extension of those presented in Cunha and Heckman (2008), with adjustments made to include the error terms for the additional measurement equations resulting from multiple types of parental investment. The model exposition presented by Cunha and Heckman (2008) contains a complete proof of the specifics of identifying this model (p. 747) and detailed explanations of the necessary assumptions discussed above. As the only difference from their original proof is that the present analysis includes additional assumptions to account for multiple latent parental investment factors, the full proof is not included in this thesis. Finally, as it is not necessary for identification, no assumptions are made about the normality of the distribution of the error terms.

3.3 Defining Parental Investment

This section outlines the strategy used in this thesis to define the multidimensional vector of parental investment included in the model of skill formation. Although it is presented within the context of the empirical applications of this thesis, this approach to identifying parental investment can be thought of as a general methodology, applicable to a variety of research questions and relevant longitudinal data.

This methodology builds on the literature discussed in Chapter 2 and synthesises several existing strategies used to define investment. I have already established that differentiating among types of investment is an important aspect of the empirical model, but specifying these different types of investment requires careful consideration of both the theoretical understanding of childhood skill development and the statistical features of the available data.

In the text that follows, I present a three-step approach to defining parental input that I will use in the empirical applications presented in Chapter 4 and Chapter 5. The first step identifies a set of relevant measures of parental input; the second step determines how many latent factors there are underlying this set of measures; and the final step assigns a factor structure. The three-steps are explored in detail below.

3.3.1 Indicator Selection: Relevant Measures of Parental Input

The first step in identifying the types of parental inputs to be included in the model is *indicator selection*, the process of choosing a set of relevant measures. As discussed in Chapter 2, there are multiple parenting factors that have been shown to predict skill development. The methodology outlined in this thesis is not intended to specifically capture all of these factors but instead lays out a methodology that can be applied to measuring the role played by various types of parental input.

Relevant measures of parental input can take a variety of forms, but in each empirical application, I identify a set of indicators using the following to two considerations:

1. ***Identifying an aspect of parenting to focus on***.

   The focus of a given research project is generally driven by the desire to investigate a specific area of parenting. Once a particular 'area of interest' is identified, indicators must be selected that are thought to capture this particular construct (or set of constructs). It is not suitable to take all possible measures of parenting and conduct an exploratory factor analysis (EFA) to see what factors are defined. Instead, there must be some pre-existing hypothesis about the general area that is to be measured. Researchers must identify a potential area of investment and explore whether this is a unidimensional or multidimensional input. For example, if researchers wanted to examine the impact of parental attitudes, they would use a set of indicators to see if attitudes are indicators of a single underlying factor (e.g. attitudes towards parenting) or instead represent several factors (e.g. attitudes towards education, attitudes towards discipline, etc.)

2. ***Data containing a sufficient number of indicators for this aspect.*** The specific parental investment indicators I select for each empirical chapter is influenced by the availability of data containing a large enough set of measures to capture the construct identified. Not only must the indicator questions be identified using existing theoretical and empirical work, but they must also be available within the data. For example, if I were interested in the impact of different discipline styles, but the data only contained one question about discipline, it would not be possible to model this construct using the present framework.

Potential Types of Investment to Focus On

The 'area of interest' can take multiple forms and the area chosen does not impact the approach to modelling. Chapter 2 laid out many different examples of parental inputs that have been shown to influence skill development, and, for the sake of brevity, a full discussion of investment types will not be repeated in this chapter. For now, it is sufficient to say that as long as a suitable number of indicators are available, the parental input vector can be used to examine a wide variety of different investment types. In order to have the necessary theoretical justification for choosing a particular research focus, I look to the existing literature on parenting and child development for guidance. The relevant details from this literature will be revisited in the empirical applications that follow. Potential areas of interest include, but are not limited to: types of household resources; parental time inputs; parenting styles; parenting behaviours; and parental attitudes. Although the full set of parental inputs will be justified by existing theoretical frameworks, the identification of the relevant set of indicators must be repeated for each construct included in the structural model. If multiple constructs are identified, they are included in the full model of skill formation using separate vectors.[68]

Modelling Financial Resources as Distinct Type of Input

As discussed in Chapter 2, family resources play a distinct role in childhood development and must be controlled for independently of other parental inputs. For this reason, I include any indicators that are driven solely by financial resources separately in the model. Including measures of SES as indicators alongside parenting behaviours assumes that SES and parenting are both aspects of the same latent factor, but this is

[68] Although this chapter focuses on identifying latent parenting factors that contribute to development, observable parent characteristics and factors that influence child development (e.g. parental education) will also be included directly in the model as controls.

difficult to justify as the financial contraints which limit financial inputs differ from the emotional and time contraints which determine parenting behaviour.

While SES should not be considered an indicator of the parenting factors being studied, it must still be included in the model. Measures of SES can either be included directly in the model as observed variables or modelled as a distinct latent factor in the full model. If, instead, measures of SES are used as indicators for the general parental inputs, the model is not able to distinguish between the impact of SES and the effect of specific parenting behaviours.

Compiling a Set of Indicators for Parental Input

There are two possible approaches when selecting indicators of parental investment. In a ‘question-driven’ empirical research project, the study begins with a pre-selected parenting construct, uses existing literature to identify indicators for this construct and then identifies longitudinal data which contains these types of indicators. Alternatively, for a ‘data-driven’ project, the research starts with a specific longitudinal data set, and finds a construct in the literature which aligns with the measures included in the data. This is not to suggest that a data-driven method would include all measures of parenting contained in the data set, but instead that it would examine the available measures and select those that appear to measure a specific aspect of parenting. With either approach, it is important to make sure that each indicator aligns with a specific parenting construct.

This thesis relies on a data-driven approach. Such an approach allows me to apply the model to two well-known longitudinal studies. With this approach, many potential indicators are in the form of existing, pre-validated, parenting assessment tools. When this happens, I take special care to understand what construct the questions were originally designed to measure. These surveys also contain potential indicators that are not part of an existing assessment tool and are simply stand-alone questions. When identifying the factor structure, I make sure to consider parenting constructs that contain both types of measures.

Once a comprehensive list of potential indicators from these surveys is compiled, the next step is to check that the relevant measures have responses that are distributed across the range of possible responses and are not so extreme as to have the majority of parents choosing the same response. For example, the question “How much do you agree with the following statement: it important to feed my child?” would in all likelihood provide very little information as almost all parents would answer “Strongly Agree.”

3.3.2 Dimensionality: Setting the Number of Factors

Once the relevant indicators of parental investment have been identified, the next step is to decide how many latent parenting factors underlie these observable measures.[69]

Existing Approaches:

Within the existing literature on parenting, defining the types of investment in empirical models is generally accomplished using one of the two approaches listed below.

1. ***Using theory:*** The first approach builds on existing research and uses accepted theories about parental input to define how many latent factors make up a given parenting construct. The indicators are manually assigned to each type of input and the model is specified accordingly. This approach fails to capture how these variables are distributed in the data being used for the analysis. If the theory used to define the number of factors and assign indicators to a specific parenting factor is inappropriate for the given data, then the structure may be misspecified.

2. ***Using statistical analysis:*** Instead of having a pre-determined categorisation of the parenting inputs, this approach uses statistical analysis in the form of principal components analysis (PCA) or exploratory factor analysis (EFA) to decide how many latent factors are captured by the set of parenting indicators. While it is able to capture the distribution of the indicators in the given data, the statistical approach is limited as there is substantial disagreement on the best strategies to determine dimensionality. In addition to the lack of consensus on methodology, measurement error, skewed indicator scales or other variations in the data are all known to result in statistical strategies providing false measures of dimensionality.

Combining these Approaches:

This thesis uses EFA to assess the dimensionality of parenting constructs in the data, and supplements the statistical results with information from the relevant literature on parenting. While it is particularly important to employ theoretical knowledge in cases where the measures of fit estimated using EFA suffer from distributional bias, it is also important to confirm any structure suggested by EFA with a theoretical explanation.[70]

Below, I review the basics of the statistical approach for determining dimensionality. Following this, I discuss the specific fit statistics which I use to determine the dimensionality of the parental construct — this includes a discussion of the limitations of EFA. This section concludes with an explantion of how supplementing the results from EFA with the theoretical knowledge explored in Chapter 2 can help address these limitations.

[69] This is often referred to as defining dimensionality or factor retention.

[70] This approach is not particularly novel, and is suggested by all proponents of EFA. Unfortunately, many studies rely on fit-statistics and do not support their structure with suitable theoretical explanations.

Using Statistical Analysis to Determine Dimensionality

Many fields of social science have a long history of using statistical techniques to identify a structure within a large set of observed variables (examples include: Cattell, 1963; Hotelling, 1933; Pearson, 1901; Spearman, 1904). There are several statistical approaches used to accomplish this task, with different fields having a preferred methodology. In psychology, factor analysis is used to develop a set of measures that capture underlying theoretical factors (e.g. Cattell, 1963; Spearman, 1904). In economics and finance, measurement reduction strategies in the form of principal components analysis are used to collapse multiple measures into a single index (e.g. Forni, Hallin, Lippi, & Reichlin, 2000; Stock & Watson, 2002). Though the methodologies vary, in each approach, statistical measures of model fit are used to determine how many factors (or components) provide the best-fitting model. This is known as determining the dimensionality of a construct.

Factor Analysis versus Data Reduction: The two most commonly used techniques for determining dimensionality are *exploratory factor analysis* (EFA) and *principal components analysis* (PCA). These two modelling strategies are defined below.

EFA: The *common factor model* presented by Thurstone (1947) is the basis for EFA. In this model it is assumed that there is some underlying latent factor which causes each of the observable indicators and that this factor is able to explain the correlation between the observable indicators. More specifically, "each measured variable in a battery of measured variables is a linear function of one or more common factors and one unique factor" (Fabrigar, Wegener, MacCallum, & Strahan, 1999, p.275). The common factor can be extracted through factor analysis, a process which compares the variances across the observable indicators. The variance for each observable indicator can be partitioned into the *common variance*, which is caused by the underlying latent factor, and the *unique variance* that is specific to each indicator (Brown, 2015, p.11). The common variance is shared by all the indicators of a given latent factor and EFA extracts this common variance in order to estimate the underlying latent variable.

PCA: Unlike EFA, PCA does not assume the existence of latent factors, and is more "appropriately used as a data reduction technique" (Brown, 2015, p.20). PCA fits a model that is best able to account for the variance in the measures of the indicators. Fabrigar et al. (1999) describe PCA as attempting "to determine the linear combinations of the measured variables that retain as much information from the original measured variables as possible"[p.275]. Importantly, while EFA assumes the presence of unique variance, some of which is the result of measurement error, PCA assumes the indicators themselves are measured without error (Schmitt, 2011). In fields such as finance, it is possible to argue that measurement error of indicators such as asset prices might be minimal, but such an assumption would be more difficult to justify in behavioural studies that use self-reported Likert-type scales to capture indicators.

Comparing EFA and PCA: Figure 3.4 provides a graphical representation of EFA and PCA. Though the models are very similar, the diagram highlights two key differences. First, the arrows point in opposite directions in the two models: for EFA, the arrows extend from the latent factor to the observed indicators; but for PCA, the arrows extend from the observed indicators towards the principal component. Second, in the EFA model it is assumed that each indicator is measured with error, compared to the PCA model where the only error is assumed to be on the principal component itself.

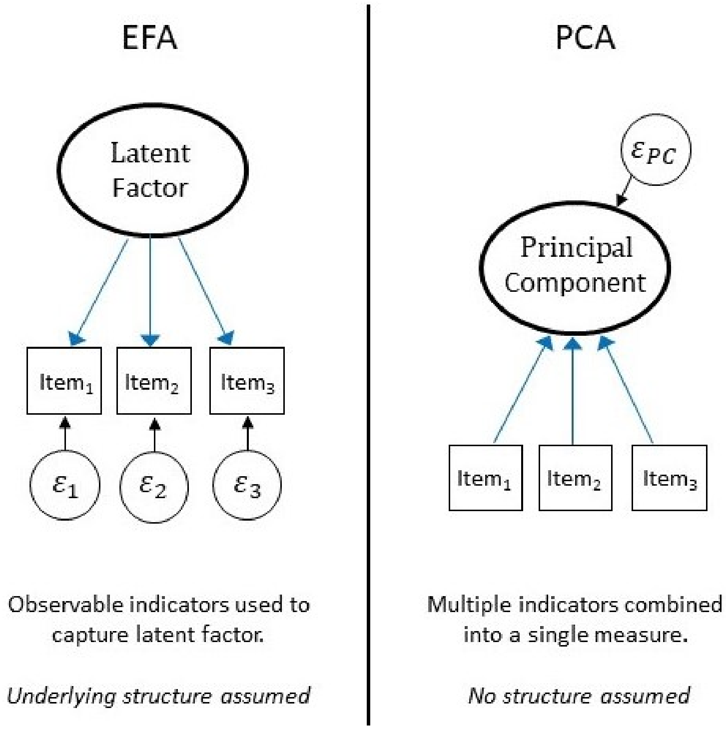


Fig. 3.4 Comparing Exploratory Factor Analysis and Principal Components Analysis

Though EFA and PCA are fundamentally different concepts, PCA is often used as an extraction technique for papers describing factor analysis.[71] Osborne, Costello, and Kellow (2008) reviewed the literature containing the keyword "factor analysis" in the methodological description, and found that 64% of researchers use PCA as an extraction technique. This overuse of PCA stems from historical limitations of statistical analysis software. As detailed in the review of Osborne et al. (2008): "PCA became common decades ago when computers were slow and expensive to use; it was a quicker, cheaper alternative to [the more computationally complex] factor analysis" (p.4). Fortunately, statistical software and computational power have drastically improved and EFA models can now be easily estimated using the appropriate methodology.

[71]Under certain conditions the results obtained from PCA can approximate those obtained using EFA (Fabrigar et al., 1999). The literature on the choice of methods is mixed, but Fabrigar et al. (1999) have presented a critical review of the literature and argued that given the recent increases in computational power, it is generally more appropriate to use EFA.

Present Context: As this thesis aims to measure latent parenting constructs and the data is assumed to be measured with error, EFA is needed to determine dimensionality. Using the set of indicator variables identified earlier in this section, it is possible to conduct EFA and compare the fit-statistics for different numbers of factors. The empirical chapters in this thesis use these fit statistics to provide initial guidance on how many factors are required to capture each parenting construct. These results are then interpreted through the lens of existing theories on child development in order to finalise the parental investment factors that I include in the structural model.

Using EFA to Determine Number of Factors: In this thesis, once the relevant indicator variables have been identified, statistical software is used to conduct an EFA. The results from this EFA can then be compared to pre-defined criteria to decide how many factors to retain for the given set of indicators (Brown, 2015). For clarity, the specific details of interpreting EFA results are reviewed below.

There are a multitude of statistical criteria (fit-statistics) that can be calculated for a given EFA, but there is no clear consensus on which of these measures is best able to determine the underlying factor structure. Fabrigar et al. (1999) provide extensive recommendations on the use of EFA in applied research and "suggest that researchers rely on multiple criteria when deciding on the appropriate number of factors to include in a model" [p.283]. More specifically Fabrigar et al. (1999) suggest the use of several descriptive fit indices[72] along with eigenvalue factor selection procedures, but note that "even the best procedures are not infallible" and EFA results should always be considered alongside "relevant theory and previous research when determining the appropriate number of factors to retain"[p.281].

The debate about the adequacy of fit-indices for factor retention in EFA is ongoing with Garrido, Abad, and Ponsoda (2016), Barendse, Oort, and Timmerman (2015) and Yang and Xia (2015) using simulated data to provide the most up-to-date assessment of the suitability of fit-indices. In line the methodological recommendations from Fabrigar et al. (1999) and Brown (2015), my proposed methodology makes use of multiple fit-statistics along with reporting the results from eigenvalue-based procedures when determining dimensionality. Each of the relevant fit-statistics and eigenvalue-based procedures are discussed below. This discussion includes a review of the relevant cutoff value or retention strategy that are used when determining dimensionality in the empirical studies presented in this thesis.

[72] Fabrigar et al. (1999) suggest chi-squared measures of fit such as RMSEA and comparative fit indices such as CFI and TLI.

- **Eigenvalue-Based Procedures for Determining Dimensionality**:

  Eigenvalues are computed from the correlation matrix and can be loosely thought of as representing the amount of variance in the indicators which is accounted for by a given factor. The larger the eigenvalue, the more explanatory power a factor has for the observed variance in the indicators. When determining dimensionality based on eigenvalues, there are three common approaches: the Kaiser-Guttman criterion, the scree test, and parallel analysis.

  Kaiser-Guttman criterion: Also known as the *eigenvalue-greater-than-one rule*, the Kaiser-Guttman criterion recommends determining how many of the eigenvalues are greater than one and then using this as the number of factors to retain (Guttman, 1954; Kaiser, 1960). This is thought to be the most commonly used factor-retention strategy: Osborne et al. (2008) surveyed over 1,700 studies involving the use of EFA to find that 45% of them rely on the Kaiser-Guttman criterion to determine the number of factors.

  Though it is widely used, there is substantial evidence that the Kaiser-Guttman criterion is prone to over-factoring (Brown, 2015; Fabrigar et al., 1999; Osborne et al., 2008). In cases where the eigenvalues fall just above or below the cutoff of 1.00, the criterion is criticised for arbitrarily defining eigenvalues just above one as factors worth retaining, but eigenvalues just below one as unimportant (i.e. 0.99 would not be retained, while 1.01 would). In this thesis, the *eigenvalue-greater-than-one rule* is applied, with special examination of any factors near the boundary. In these boundary cases, any decisions on dimensionality rely on logic and other indices.

  Scree Test: The second-most popular factor retention strategy is Cattell's (1966) scree test, which is based on a visual interpretation of a plot of eigenvalues. To conduct a scree test, a graph constructed with the eigenvalues on the vertical axis and the corresponding number of factors on the horizontal axis. The number of factors retained is based on a visual inspection of this graph to determine where the last substantial vertical drop occurs.[73] This is referred to as the elbow or inflection point. The number of retained factors is set equal to the number of eigenvalues before this drop-off. In some cases this drop-off is very clear; in others, however, the scree test approach is rather subjective as there is no standard definition of a 'substantial drop-off.' Even with this limitation, the scree test is still a popular approach.[74] As with the Kaiser-Guttman criterion, this factor retention strategy has the potential to be quite arbitrary and requires special examination of the factors near the boundary. Therefore, this thesis only considers visual interpretation of scree plots alongside other measures of dimensionality.

[73] This is alternatively described as the point after which the eigenvalues are relatively stable.

[74] A literature review conducted by Osborne et al. (2008) found that scree tests are employed by 42% of researchers using factor analysis.

Parallel Analysis: Parallel analysis was introduced by Horn (1965) for use with PCA. It compares the eigenvalues from the observed data with a corresponding set of eigenvalues generated from multiple datasets which contain uncorrelated indicators. These datasets are generated to have the same number of indicators and sample size as the observed data, but the values are randomly generated. The number of factors retained corresponds with the number of observed eigenvalues that are greater than their corresponding random eigenvalues.

Fabrigar et al. (1999) argued that parallel analysis is the gold-standard method of factor retention. Although Garrido et al. (2016) used simulated data to demonstrate how parallel analysis can be applied to categorical data, this methodology is not available in major statistical software packages (Muthén & Muthén, 2017). While it is possible to obtain parallel analysis results by treating the indicator variables as continuous, there is insufficient evidence that continuous models capture the properties of categorical variables. For this reason, this thesis does not use parallel analysis.

Figure 3.5 provides an example of plotted eigenvalues from observed data[75], along with eigenvalues generated from parallel analysis and a horizontal line for eigenvalues equal to one. This example is illustrative of how the three eigenvalue-based factor retention methods can yield different results. Using the Kaiser-Guttman criterion, there are four factors that are selected. The scree-test yields for this particular example can be interpreted in several different ways. It suggests a 3-factor model or 6-factor model depending on the subjective interpretation of where the eigenvalues level off. Finally, the parallel analysis depicted in this example would suggest the use of a three-factor model.

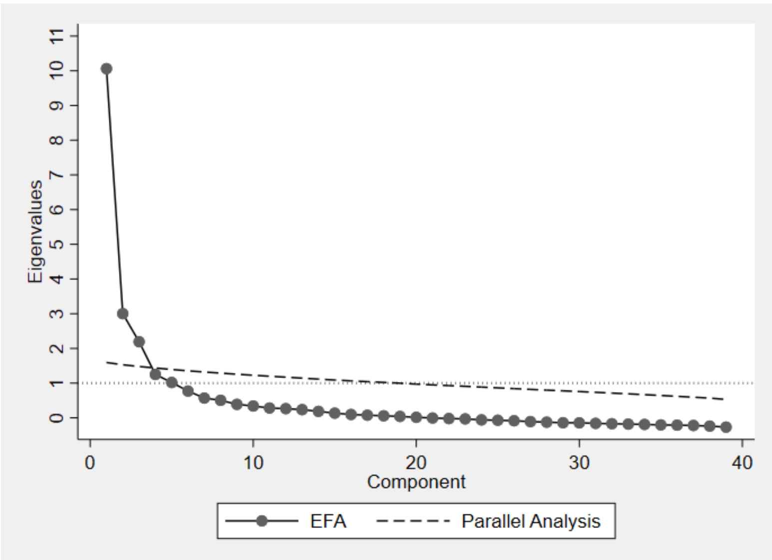


Fig. 3.5 Eigenvalue-Based Factor Retention Methods

[75]Data used for this example are unrelated to the analysis presented later in this thesis and are only used to illustrate the concept.

- **Goodness-of-Fit Statistics and Dimensionality**:

  Due to the limitations of the eigenvalue-based approaches discussed above, methodologists recommend using them along with measures of model fit (Barendse et al., 2015; Fabrigar & Wegener, 2012; Fabrigar et al., 1999; Yang & Xia, 2015). These measures of model fit are known as *fit-indices* and were originally developed for structural equation modelling (SEM) and confirmatory factor analysis (CFA). Fit-indices are a measure of how well a given model is able to approximate the observed data. The estimated fit-indices can be compared to pre-defined cutoffs that have been established as providing an adequate level of model fit.[76]

  To determine the number of parenting factors, I begin by estimating a one-factor model and calculating the fit-statistics. Repeating this process for a sequence of models, each with one additional factor, allows me to compare the fit-statistics of the differently factored models. The number of factors that I retain corresponds with the model that satisfies the accepted cutoffs and fits substantially better than a model with one less factor.

  Unfortunately, the relevant cutoffs were established by evaluating how well CFA models fit simulated data, and there is limited evidence on how suitable these cutoffs are for determining dimensionality through EFA (Fabrigar & Wegener, 2012). Recent research using simulated data has shown that the CFA based cutoffs might be insufficient for correctly determining the number of factors in skewed categorical data, but can provide reasonable accuracy for large samples of unskewed data (Garrido et al., 2016; Yang & Xia, 2015). The relevant fit-indices are presented below, along with details about specific considerations for their use in EFA.

  ***RMSEA*:** The Steiger-Lind root mean squared error of approximation (RMSEA; Steiger, 1990, 2016) is an absolute fit index which estimates the unexplained variance, thus comparing the proposed model to a model that perfectly captures the data. RMSEA is reported using values between 0 and 1, with larger values indicating poorer model fit. For CFA, an RMSEA <0.05 indicates a close fit, while <0.08 indicates a reasonable fit (Browne & Cudeck, 1992; Hu & Bentler, 1998).

  Conducting EFA on simulated data, separate studies by Garrido et al. (2016) and Yang and Xia (2015) found that for categorical variables, RMSEA is relatively accurate given suitably large sample size, and relatively unskewed distribution of the responses.[77] Garrido et al. (2016) recommended using lower cutoff values for RMSEA (ranging from 0.03-0.05) to determine dimensionality from EFA [p.9]. This thesis uses 0.05 as an absolute cutoff for reasonable fit.

[76] For further details on the use of fit-indices in CFA, the reader is directed to Browne and Cudeck (1992) and Hu and Bentler (1999).

[77] The largest samples analysed by Yang and Xia (2015), and Garrido et al. (2016), were 400 and, 1,000 respectively.

SRMR: Another absolute fit measure, the standardised root mean-square residual (SRMR; Bentler, 1995; Jöreskog & Sörbom, 1982), measures the standardised difference between the predicted and observed correlations. SRMR values fall between 0 and 1, with 0 representing a perfectly fitting model. For CFA, SRMR<0.05 indicates a good model fit, while SRMR<0.08 indicates a reasonable fit (Hu & Bentler, 1998).

When using SRMR to determine the number of factors in simulated data, both Garrido et al. (2016) and Barendse et al. (2015) find that lower cutoff values tend to be too stringent for EFA and result in over-factoring. Garrido et al. (2016) find that SRMR is highly inaccurate for identifying the number of factors categorical data and recommend the use of other fit-indices where possible. Although SRMR is not used for the basis of any dimensionality decisions, it is standard practice to report this value, so I include the estimated values of SRMR in my results.[78]

CFI**:** The Bentler (1990) Comparative Fit Index (CFI) is an incremental fit index which compares the proposed model to a baseline model where all variables are uncorrelated. CFI is reported as a value between 0 and 1 with higher values indicating a better-fitting model compared to the baseline model which corresponds with a CFI of 0. The cutoff values for CFI are discussed below, alongside the TLI cutoff values.

TLI**:** The Tucker-Lewis Index (TLI; Bentler & Bonett, 1980; Tucker & Lewis, 1973), is another incremental fit index. The TLI contains an adjustment to compensate for the additional degrees of freedom resulting from increased model complexity and to penalise for additional variables that do not improve model fit (Brown, 2015). Values of the TLI generally fall between 0 and 1, but, unlike the CFI, the TLI is not normed, so TLI scores may fall outside this range. These outliers aside, the TLI is interpreted in the same way as the CFI, with higher values indicating a better-fitting model.

For CFA, values of CFI/TLI > 0.95 indicate good fit while "values in the range of .90 — .95 may be indicative of acceptable model fit" (Brown, 2015, p.75). This range is based on the works of: Hu and Bentler (1999) which argued that "a cutoff value close to 0.95"(p.27) provides good fit for continuous variables; and Bentler (1990) which indicated that a cutoff of 0.90 provides acceptable fit in certain situations.

For EFA, Garrido et al. (2016) found that for categorical data, CFI and TLI perform nearly identically and have a high level of accuracy when used to determine the number of factors. However, for skewed categorical data, Garrido et al. (2016) found that reducing the cutoff value to 0.90 reduces error when determining dimensionality, as there is "notable bias toward over-factoring across the [traditional] cutoff values"(p.16). In light of these findings, the empirical chapters of this thesis use 0.90 as an absolute cutoff and interpret values from 0.90 to 0.95 alongside other measures of fit.

[78]It is accepted practice to follow the guidelines for CFA presented by Hu and Bentler (1998). These guidelines recommend reporting SRMR along with one other index.

Interpreting EFA Results

It can be seen from the information presented above that determining the number of factors using EFA and applying rules of thumb may not always correctly identify the number of factors in the data. Though fit-statistics provide useful information to guide decisions in regards to dimensionality, there is substantial evidence that fit statistics do not always accurately identify the number of factors contained in simulated data (Barendse et al., 2015; Garrido et al., 2016; Yang & Xia, 2015; Yu, 2002).

To provide a sense of how inconsistent these measures are, it is helpful to briefly review a study by Garrido et al. (2016) which evaluated the performance of the four fit indices described above and compared them to the results obtained using parallel analysis. The authors used Monte Carlo methods to generate multiple data sets with variables that load onto a given number of factors. They then used EFA and applied various measures of fit to determine the number of factors the model estimates for the given data. This estimated number of factors was then compared to the number of factors that were used to create the simulated data. This methodology allowed the researchers to see how accurate each of the estimators were at determining dimensionality under a variety of conditions.

Using this methodology, Garrido et al. (2016) found that none of the strategies discussed above are able to predict the correct number of factors more than 90% of the time, and in many cases the strategies are far less successful. Though the models were tested across a variety of sample sizes and with various factors, it is simplest to examine the maximum level of correct identification across all conditions. A summary of this 'maximum accuracy' of the various fit statistics is presented in Table 3.1.

Table 3.1 Percentage of Cases where Given Fit Statistic Correctly Identifies the Number of Factors

	Parallel Analysis	RMSEA	CFI	TLI	SRMR
Unskewed Categorical Variables	86%	57%	80%	79%	57%
Skewed Categorical Variables	78%	59%	69%	67%	45%

Source: Adapted from the results presented by Garrido et al. (2016).
Notes: RMSEA = root mean square error of approximation. CFI = comparative fit index. TLI = Tucker-Lewis index. SRMR = standardised root mean square residual.

The values presented in Table 3.1 represent the best case situation in terms of correct estimates of dimensionality. These high levels of accuracy generally occur when EFA is conducted on larger sample sizes, for data that has low levels of correlation between the factors and with responses spread across multiple response categories. The detailed findings of Garrido et al. (2016) show that, in less ideal conditions, the accuracy of the fit statistics in identifying the number of factors is even lower.

The results from Monte Carlo simulation studies such as Garrido et al. (2016) and Yang and Xia (2015) demonstrate that even when data is created using pre-existing factors, EFA may incorrectly estimate the number of factors. What these studies do not show is that, when applied to observational data, EFA may also: identify factors that have limited conceptual basis; or, due to biased data, fail to identify a known construct.

For this reason, I follow the recommendation of Brown (2015), who argued that EFA should simply be thought of as a way to "identify the solution that reproduces the observed correlations considerably better than ... models involving fewer factors" [p.26]. In the context of this thesis, this means that the set of parental inputs identified by statistical analysis as producing the best-fitting model, might not be the set of factors that are best-suited to studying the corresponding theoretical parenting construct.

Ensuring that Estimates of Dimensionality are Conceptually Sound

The inability of EFA to consistently identify the number of factors, and the need for factors that can be readily interpreted indicates the importance of using EFA results as part of a larger strategy for determining factor dimensionality. When discussing the use of EFA in applied research, Henson and Roberts (2006) suggest using "both multiple criteria and reasoned reflection" when determining dimensionality [p.399]. Though some existing research on parental inputs and parenting style relies on EFA or PCA alone, the role of theory in interpreting EFA results is generally accepted by methodologists, with many methodological texts recommend using EFA as a means to guide a conceptually realistic interpretation of the data (Brown, 2015; Fabrigar & Wegener, 2012; Kline, 2015).

Using existing knowledge from developmental psychology, economics and education research, I am able to examine the estimates of dimensionality provided by EFA to decide how these estimates should inform my structural model. To do this, each empirical analysis in this thesis begins by using EFA to estimate the number of parenting factors and then examines the rotated factor loadings provided by EFA to assess the conceptual validity of each suggested parenting factor. An item is assumed to contribute meaningfully to a given factor if it has a rotated loading of above 0.3 for that factor (Brown, 2015). If the set of 'meaningful' indicators for a given factor can be readily interpreted using a construct from existing theories on parenting of or child development, then the estimate of dimensionality suggested by EFA is used in my model.

In some cases, EFA results may suggest that the number of factors could take several possible values (i.e. different fit-indices suggesting conflicting numbers of factors). In such ambiguous cases, factor loadings are examined for each of the possible estimates of dimensionality and the number of factors retained corresponds with the number which has the strongest conceptual base. In cases where factor structures cannot be conceptually explained, I reconsider the choice of indicator variables and examine the factor structure that results from removing poorly fitting indicators.

The way in which each empirical application uses theory to help determine dimensionality depends on: the type of indicators used, the results of the factor analysis, and the particular research question being asked. Though the specific details vary, I address each of the following three questions when determining the number of parenting factors to include in each empirical application of the model.

1. *Can each of the factors identified in the model be easily interpreted using existing theoretical constructs or logical examination of the loaded indicators?*

   The parenting factors identified in this thesis are not all based on strictly defined theoretical constructs, but each of them has some logical explanation that can be situated within the literature. For example, if existing theory proposes that 'parent-child-interactions' are a type of parenting factor, but the results from EFA may suggest loading artistic parent-child activities separately from athletic activities, my model would separate parent-child-interactions into two different factors based on type of activity (in this case athletic and artistic.) This approach allows me to use the factors suggested by EFA to identify sub-categories of known parenting constructs that are based on measured variance within the data.

2. *Do the results from EFA suggest that this is the best factor structure?*

   In the empirical chapters, the number of factors suggested by EFA is not always clear-cut. In ambiguous cases, there may be several models which provide acceptable fit or different fit-indices suggesting conflicting numbers of factors. When this happens, I report the values for each measure-of-fit and briefly discuss why the theoretical basis for a given parenting construct has led me to prioritise the dimensionality suggested by one measure of fit over that suggested by another of the fit-indices.

   Similarly, my interpretation of EFA results always considers the type of indicator measures used and the distribution of the responses in the given data. For example, when the indicators used have a skewed distribution of responses, I apply more stringent cutoff values to avoid over-factoring.

3. *If the chosen factor structure is not the best-fitting model, do the fit-indices suggest reasonable fit?*

   In this thesis, there are cases where the number of factors included in the full model does not correspond with the best-fitting model obtained using EFA. Often, this involves the type of boundary cases discussed above, but for any model, I take special care to discuss the measures of fit for the final factor structure.

3.3.3 Assigning the Factor Structure

The final step for modelling parental investment is to assign each of the indicator variables to one of the identified factors. This is done by applying factor rotation and choosing a solution which assigns indicators to specific factors. This is known as *identifying the factor solution*.[79] This process is often thought of as part of EFA, but for this thesis the parenting constructs are a key component of the larger model. Therefore, in the text that follows, I specify exactly how the factor structure is determined.

Accepted Best Practices for Identifying a Factor Solution

When assigning the factor structure, the goal is to identify a set of factors that is readily interpretable and can be explained using existing theory. Brown (2015) outlines the accepted best practice for using EFA results to define a suitable factor solution and identifies the following four criteria that must be satisfied by any factor solution.

- Indicators should only be retained if they have a loading that is at least 0.3.[80]
- The indicators that are assigned to each factor should be interpretable as having a common theme or underlying construct.
- The assigned factor structure must satisfy the five criteria of *simple structure* defined by Thurstone (1947). These are as follows:
    1. Each indicator variable should have a loading of zero for at least one factor.
    2. Each factor should have at least m loadings at or near zero, where m is the number of factors in the model.
    3. Each pair of factors should have several indicators that have zero loadings for one factor but not for the other.
    4. In models with more than four factors, each pair of factors should have several indicators that have zero-loadings on both factors.
    5. A pair of factors should have only a few indicators loading onto both factors.
- The factor rotation selected must be justified by an understanding of the data. In this thesis, I use an oblique factor rotation. This is because parental input factors are often correlated and oblique rotations allow factors to inter-correlate.

[79] A review of factor rotation and selecting a factor solution is beyond the scope of this thesis. For further details, the reader is directed to Browne (2001) and Fabrigar and Wegener (2012).

[80] In some cases indicators might be retained slightly below this cutoff if there is strong theoretical justification for their inclusion in the factor.

Additional Considerations for this Thesis

In addition to satisfying the general criteria listed above, the definition of parenting used in this thesis also considers how the factor solution fits within the larger structural model. To accomplish this, I make two additional considerations when identifying the factor solution. These are: consistency across periods in the model, and, in certain cases, the need for agreement with factor structures used to create the data.

Consistency Across Observational Periods: Due to the longitudinal nature of the model of skill formation, the same measures will often be repeated across multiple periods. In cases with repeated measures, I conduct EFA separately for each period of observation. If the results from EFA suggest slightly different factor solutions across periods, I consider assigning a consistent factor structure across periods. This is not to say that ill-fitting models are accepted for the sake of consistency, but instead to suggest that in ambiguous cases a consistent structure might be the deciding factor for choosing a particular factor solution. For example, if a given indicator loads strongly onto a theoretically justified factor in most periods but is only weakly loaded in a single period, it might be worth retaining this weakly loaded indicator so the impact of the construct can be consistently discussed over time. In cases such as these, my empirical results will review the fit-statistics of models which use both factor solutions and provide justification for my final modelling decision in the presentation of the final model.

Agreement with Factors from Survey Design: Often, the relevant secondary data will include measures of parental investment that correspond with pre-existing factors. In such cases, I consider these pre-existing factor structures when specifying my model. As with all other considerations, the final structure will be decided after examining the relevant results from factor analysis and considering the goals of the research project. For example, if a cohort study has used a pre-validated set of parenting measures that are designed to measure a set of factors, it is likely that the results from EFA will capture the pre-designed factor structure. In cases where the factors identified using EFA are inconsistent with the factors defined by the survey, I consider both the pre-validated structure from the survey, alongside the data-driven factor structure and provide careful justification for the factors I include in my final model.

Resulting Factor Solution – Parental Investment

Combining these thesis-specific considerations with the standard approach for identifying a factor solution allows me to define a factor structure for parental investment that is both data-driven and theoretically justified. As this procedure requires some subjectivity, the empirical applications in this thesis each provide a detailed explanation of how I assign the indicator measures to specific parental investment factors.

3.4 Empirically Estimating the Model

The previous sections have outlined the procedures and modelling considerations for specifying the full model of skill formation. This section explores the methodological details of empirically estimating this model using existing data. The relevant methodological considerations for estimating the model can be considered to fall into three areas: data requirements, statistical considerations and analytical procedures. While many aspects of estimating this model are specific to each empirical project, several methodological elements concern all empirical applications of the skill formation model. These general methodological concerns are reviewed below, while study specific details are presented later in this thesis in Chapter 4 and Chapter 5.

Data Requirements

The empirical chapters in this thesis make use of existing secondary data. Although this approach is common in other disciplines, readers from certain sub-fields of education may be unfamiliar with some of the considerations made when selecting appropriate data for this type of analysis. Just as a purposive study would pre-define the target sample size, appropriate time-frame of data-collection and suitable measures, I identified what the adequate sample size, sufficient time-frame and suitable measures were for estimating the model, before selecting the secondary data to use in my empirical chapters. Details of these three considerations are outlined below.

Sample Size

To determine the required sample size for my empirical applications, I turn to the literature on structural equation modelling. Opinions vary on how to define the appropriate sample size for SEM, with approaches falling into two general categories. The first type of approach advises the use of pre-defined 'rules of thumb', while the second type of approach recommends using a power analysis which calculates an adequate sample size based on model characteristics such as: the number of indicator variables; the number of latent variables; and the desired statistical power.

Researchers who use the first approach define sample size using minimum cutoffs for either absolute sample size or for the ratio of parameters to observations. This approach is the most widely used but Osborne et al. (2008) have shown that the exact 'rules' applied vary substantially across studies. For example, Hair (2018) proposed that, for complex models with multiple latent constructs, a minimum sample size of 500 is required for model stability. Alternatively, Kline (2015) advised an absolute minimum sample size of 200, and recommended that researchers aim to satisfy rules regarding the minimum ratio between sample size and number of parameters (N:q) that were identified by Jackson

(2003). The rules presented by Jackson (2003) suggest that the sample size to parameter ratio should be at least 5:1, and ideally 20:1. For this thesis, I avoid using ratio-based rules because studies by Wolf, Harrington, Clark, and Miller (2013) and MacCallum, Widaman, Zhang, and Hong (1999) have shown that these rules are inappropriate for more complex models with large sample sizes.

The second type of approach uses existing software or specialised estimation tools to conduct a power analysis and identify the required sample size. This approach builds on recent research which uses simulated data to test the validity of previously accepted sample size rules, and finds they do not always work. In response to this growing evidence base, Brown (2015) argued against "rely[ing] on general rules of thumb that seem to persist in the applied research literature (e.g. minimum sample size)" [p.395], and proposed that, instead, characteristics of the model should be used to estimate the adequate sample size. Similarly, In'nami and Rie Koizumi (2013) contended that best practice in SEM should involve "using more empirically grounded, individual-model-focused approaches to determining sample size in relation to parameter precision and power"[p.39]. Readers who are interested in the specifics of this approach, are directed to Kline (2015, p.290-292) or Kelloway (2015) — both provide excellent reviews of these tools and include information on how to estimate relevant sample sizes in practice.

This thesis uses statistical tools to determine the sufficient sample size for estimating the models which I specify. The details of how this methodology is applied are discussed in each empirical application. For now, although the exact sample size will vary between empirical applications, the literature would suggest defining a starting point for my required sample size by applying the rules of thumb regarding minimum sample size for SEM. Specifically, identifying a sample of at least 500 individuals will exceed the minimum sample size recommendations of both Kline (2015) and Hair (2018).

Longitudinal Coverage and Relevant Data

To estimate the model, I need data which follows the same children over multiple years and contains relevant measures of cognitive ability, non-cognitive ability, parental investment measures and demographic measures.[81] Secondary data in the form of *cohort studies*, is perfectly suited for my empirical chapters as these studies not only follow children throughout childhood, but they also contain sufficient information on the relevant measures.[82]

[81] The details for each of these types of measures have already been presented earlier in this chapter.

[82] A cohort study follows a group of individuals over time and collects information from this group at multiple intervals (Cohen, Manion, & Morrison, 2007). The longitudinal nature of cohort studies allows for a "prolonged study of the lives of one group of respondents" (Gorard, 2001, p.86), which is ideal for research on child development. The data from cohort studies is generally of high quality as the studies are subject to significant oversight from the government agencies which commission them. Additionally, cohort studies often offer large sample sizes that an individual researcher would be unable to obtain.

Statistical Considerations when using Data from Cohort Studies

The cohort studies analysed in the empirical chapters of this thesis employ specific sampling strategies to select a sample that can be used to provide nationally representative estimates. These sampling strategies are often ignored by researchers who are interested in examining how a model holds for a given sample — such results are valid in their own right. However, for my purposes, I am interested in results that are representative of the larger population. By applying the appropriate estimation techniques, I am able to estimate models that have robust results enabling extrapolation to the reference populations for each cohort study.

Before presenting the methodology which I use to adjust for these sampling strategies, I review the specific aspects of sample design used in cohort studies. Specifically, I summarise the relevant details of sampling strategies before discussing the importance of attrition and item-non-response in cohort studies.

Understanding Sampling Strategies

As with any discussion of sample selection, there are two aspects of a cohort study's sampling strategies that I must consider before empirically estimating my model. These are: sample size and sampling technique. Sample size requirements have already been discussed earlier in this section. Below, I review sampling techniques that are commonly used by cohort studies. For clarity, this discussion of sampling techniques includes some basic definitions that will already be familiar to most readers.

In general, sampling techniques fall into two categories: *probability* or *random sampling*[83], and *non-probability* or *purposive sampling*[84] While most researchers are well acquainted with both of these sampling techniques when used in isolation, cohort studies tend to use a sampling strategy which combines these techniques. This is known as a *modified random sample*, the specifics of which are reviewed below.

The primary motivation behind modified random sampling is to create data with sufficient richness to represent small subgroups within the population while simultaneously maintaining some of the desirable statistical properties that come with random sampling. More specifically, cohort studies are often designed to have samples that are not only representative of the larger population of children born in the reference period, but also to define a sample with large enough sub-samples of various subgroups within the population. These sub-samples need to have sufficient sample size to provide adequate statistical power when comparing the subgroup to other parts of the sample. To do this,

[83]Individuals are selected randomly from the population. This means that each individual in the population has the same probability of being chosen for the sample.

[84]Individuals are specifically chosen from the general population based on some pre-determined research need or criteria. Different individuals in the population have different probabilities of being selected for the sample.

there must be a non-zero probability that individuals in each of these groups are selected. Although simple random sampling would, by definition, allow for a non-zero probability of selection Cohen et al. (2007), the population of many ethnic and minority subgroups is so small that the sample size within a simple random sample would be unlikely to be large enough to allow for statistical analysis of these groups.

Stratification and cluster sampling are two specific sampling strategies that are commonly used in cohort studies to ensure sufficient sample size for these subgroups.

- *Stratification* is a sampling technique that divides a population into distinct groups using measurable characteristics and then takes a random sample from each of these groups. For example, a population could be divided based on gender and then a random sample be taken within each gender. By using the stratification process, researchers can ensure that an adequately sized sample is collected from each group. This differs from a simple random sample, which, by design, has a non-zero probability of creating a sample that does not include a given subgroup.

- *Cluster sampling* also involves dividing the population into a number of subgroups, but unlike stratification which samples within these groups, cluster sampling takes a simple random sample of the clusters and then collects data only from these selected groups in order to represent the whole population. This is often done when it would be logistically or financially unfeasible to use random sampling.

Item Non-Response and Attrition

One unfortunate feature of cohort studies is that the data is subject to attrition and item non-response. *Attrition* describes respondents who are part of an initial sample but do not complete the full study. This can be especially problematic in longitudinal studies as respondents might move or be difficult to contact for later sweeps. In contrast to attrition, a respondent with *item non-response* has completed a survey or interview but has not responded to one or more questions. These patterns of non-response may result from avoiding particular questions or might be due to respondent fatigue. There is substantial evidence that attrition and item non-response do not occur randomly and that they tend to occur predominantly in specific subgroups of the population. For this reason cohort studies often use strategies to attempt to adjust for these sampling issues.

Empirically Adjusting for Sample Design, Item Non-Response and Attrition

To adjust for the sample characteristics described in the previous section, I make use of two estimation strategies: sampling weights and bootstrapping. Within the literature that conducts secondary analysis of survey data, both of these techniques are fairly commonplace, but their use may be unfamiliar to researchers who primarily use experimental data. Below, I provide a very brief introduction to the use of these estimation strategies for those unfamiliar with the technique.[85]

Using Sampling Weights: A sampling weight captures the probability that an individual is sampled. Put differently, "the sampling weight of unit i can be interpreted as the number of population units represented by unit i" (Lohr, 2009, p.40). In a simple random sample all individuals in a population are equally likely to be sampled and share the same sample weight, which can be simplified to one. For surveys with more complex sampling strategies, the sample weight is calculated so that it represents the number of people within the population that the sampled individual is supposed to represent.[86]

Because the concept of sampling weights is key to my analysis, it is useful to provide an overly simplified example to illustrate its importance. Consider a survey, which selects one person from each country and measures his or her height with the intention of estimating the mean height of a human being. Using the usual formula to calculate a mean, the mean height in the sample would be calculated by summing the heights and dividing by the sample size. This could be considered an un-weighted estimate of the mean height of the population. By contrast, if sampling weights were used, the mean height of a human would be estimated by multiplying each of the observed heights, by the number of people in the corresponding individual's country, and then dividing by the total world population (assuming the population information is available). In this particular example the height of a person in India would need to be multiplied by approximately one billion, while the height of a British person in the sample would only need to be multiplied by approximately 60 million. Assuming that height is a characteristic that differs by nation and that the selected individuals are representative of the height of the people in their nation, this weighted strategy yields an estimate closer to the true population mean than simply calculating the sample mean.

Though the sampling strategy and corresponding estimation approach used in this example are much simpler than the types of sampling strategies employed by cohort studies, this example illustrates how the use of sampling weights provides statistical estimates that more closely reflect the distribution of characteristics present in the reference population. Obviously, more complex statistical strategies are needed to adjust

[85] More detail on these concepts can be found in any survey analysis textbook (e.g. Lohr, 2009).

[86] Census or other population-level data provides the population frequency of certain characteristics and this information is used to calculate sampling weights.

for the sampling strategies employed by the cohort studies used in this thesis. In each empirical chapter I explain the specific details of both the sampling strategy used by the relevant cohort study, and how sampling weights are used in the estimation.

Using Bootstrapping: Although applying the provided survey weights will yield parameter estimates that reflect the distribution of the reference population, the corresponding standard errors are not a reliable estimate of the sampling variance, and often overestimate the statistical significance of the parameters. In some cases, the complex sample design, non-response adjustments and other factors contributing to the survey weights make it impossible to calculate the sampling variance using traditional econometric methods. Since variance plays a critical role in determining statistical significance, these unadjusted estimates of variance may result in measures of significance which are incorrect. In Chapter 5, I make use of an estimation technique known as *bootstrapping* to adjust for sampling strategy to ensure the variance assigned to point-estimates are representative of the full population. The specific procedural details and the use of bootstrapping are discussed in Chapter 5, but the underlying logic is briefly outlined for readers unfamiliar with the concept.

Bootstrapping can be done using pre-provided bootstrap weights or using bootstrap weights created by the researcher. These weights represent a random sampling from within the survey sample, with weights assigned to reflect the larger survey population. Re-estimating the model of skill formation repeatedly using each of these sets of bootstrap weights provides multiple estimates for each parameter. The variance of these multiple estimates is averaged to provide a reliable estimate of the sampling variance for the estimate obtained using the original sample weight. This estimate of sampling variance allows for the calculation of test statistics which correctly establish statistical significance.

Analytical Procedures

The statistical considerations described above will require the use of statistical software that is capable of using complex sampling weights in addition to being able to model structural equation models that use multiple categorical variables. The software that is best suited to meet these requirements is MPlus 8.0 (Muthén & Muthén, 2017).

WLSMV: One key feature of MPlus is that it allows for models to be estimated using a *robust weighted least squares mean variance adjusted estimator (WLSMV)*. This estimation strategy is not included in most statistical software and is specifically designed to be used within structural equation models. WLSMV is especially well suited to the estimation of latent variables that have categorical indicators.[87]

[87] Rhemtulla et al. (2012) have provided evidence that WLSMV is the estimation technique best-suited to models with ordered categorical indicator variables.

3.5 Chapter Summary

In this chapter, I have presented a comprehensive examination of the updated skill formation model that this thesis uses to measure the role that parental investment plays in the developmental trajectories of childhood skills. Though the details examined in the previous four sections are critical to understanding the specification of the model — as well as contextualising it within the literature discussed in Chapter 2 — they provide substantially more methodological detail than would be expected for any individual empirical application of the model. Presenting these methodological elements and the accompanying review of statistical concepts, allows me to set the framework for the empirical chapters that follow and to clarify how this modification to an existing model allows me to examine child development in a new way.

By adapting three elements of an existing model, the framework outlined above allows Chapter 4 and Chapter 5 of this thesis to provide a novel application of the skill formation model to data from the UK and Canada. Specifically, I modify three aspects of how parental investment is defined within the skill formation model:

1. Define parental investment as multidimensional.
2. Combine data-driven and theoretical approaches to define these parental investment factors.
3. Specify parental inputs in such a way as to separate the impact of SES from the effect of specific parenting behaviours.

I am not the first researcher to include multiple types of parental investment. Nor am I the first to use exploratory factor analysis in order to define these investment types or to examine the distinct role that SES plays in development. However, my contribution is to combine these elements within a single model, and to complement the data-driven EFA with a careful consideration of theoretical factors underlying the types of parental investment. Once investment factors have been specified, the full model can then be estimated as described in the statistical considerations and analytical procedures discussed in Section 3.4.

CHAPTER 4

Empirical Application I: United Kingdom

Using data from the UK Millennium Cohort Study, this chapter presents an empirical application of the methodology introduced in Chapter 3. The estimates from this model offer valuable insights into the developmental trajectories of skills in British primary school children and the role that parenting behaviours play in shaping these trajectories. This chapter also serves to demonstrate how the modifications I have made to the human capital production function can be applied in practice. Although the previous chapters have provided detailed empirical and theoretical justification for this model design, this chapter serves to establish the model's suitability for analysing existing data.

The chapter is organised in the following manner. First, Section 4.1 introduces the present study, and contextualises it within existing UK research. Next, Section 4.2 describes the data that is used for this empirical application, examines the variables chosen for the analysis and provides descriptive statistics for the survey sample and relevant measures. Section 4.3 reviews the empirical model within the context of the present study and discusses the methodological considerations needed for this analysis. Section 4.4 presents the results obtained using this methodology. The final section, Section 4.5, offers concluding remarks on this empirical work and its importance for policy.

4.1 Introduction

This chapter uses existing longitudinal data from the UK Millennium Cohort Study (MCS) to provide the first empirical application of the methodology outlined in Chapter 3. This empirical application allows me to satisfy two of the primary goals of this dissertation. First, it confirms the suitability of my modified model to estimate the trajectories of skill development. Second, the estimation of the model provides detailed estimates of the role that different types of parental input play in skill formation in the UK — these estimates serve to build on the existing literature and are valuable in their own right. Before introducing the empirical study, I briefly discuss how this chapter fits within this dissertation and the existing literature.

As discussed in Chapter 2, there is a substantial body of research on skill development in children, and this literature contains a wealth of UK-specific studies. Several of these empirical studies have even used variations of the human capital production function to model skill development, but in these studies there has been relatively little focus on different types of parental input. Many of these studies combine multiple parenting indicators to form a general measure of parenting, and often measures of SES are grouped in with other types of inputs. This chapter of my dissertation focuses on using existing data to provide updated estimates of skill formation in the UK, with particular attention paid to differentiating between types of parental investment, and then adjusting the skill formation model to account for different types of investment.

This dissertation is not the first empirical application of the dynamic model of skill formation. Using US data, the works of Cunha and Heckman (2008) and Cunha et al. (2010) were able to provide evidence on the potential of such models to measure skill development empirically. The papers of Cunha and Heckman focused primarily on introducing the model, and the underlying statistical assumptions that must be dealt with when modelling the relationship empirically. While these findings are critical in demonstrating the dynamic nature of skill development, the analysis was conducted on a very narrow sample (i.e.: white males, born in the US) and the analysis was limited to the specific set of parental inputs available in the data. It is unlikely that estimates of the skill development of white American boys, and focused on a specific set of parenting inputs, are able to capture the full scope of skill development experienced by children in the UK. For this reason, it is important to extend the analysis not only to UK data, using a representative sample of the entire population, but also to explore a variety of measures for parental input. More specifically, it is crucial to distinguish between family characteristics and parental investment. Human capital theory predicts that both might play a role in skill development, and it is critical for policy-makers to understand the specific and separate roles that both parental behaviour and family background play.

This chapter seeks therefore to provide updated evidence on the relationship between parenting behaviour and child skills by applying the skill formation model to a modern sample of children in the UK. The educational context in the US compared to the UK is very different; as such, it is possible that the relationships observed between parental inputs and child outcomes may also be very different in the two countries. In both countries, poorer children perform less well in school. In the US, school funding is based on local income and property taxes, so children from poorer families tend to attend schools with far lower levels of funding — and hence these schools are often of lower quality. In the UK, funding is, to some degree, compensatory; funding is deliberately focused on pupils and indeed schools in more disadvantaged circumstances. This certainly does not, however, sever the link between family background and pupil achievement in the UK, but it may mean that there is a different relationship between parental inputs in the US and the UK. For these reasons, evidence from the UK context on the relationship between parental background and pupil achievement is essential.

The analysis presented in this chapter relates to a recent paper by Hernández-Alava and Popli (2017) who were the first to apply the skill formation model to data from the UK. They applied a modified version of the skill formation model to a subsample from the same dataset used in this dissertation — namely the UK's Millennium Cohort Study (MCS). Though their analysis provided insight into skill formation in the UK context, the scope of their research was limited as it only included data from birth to age 7 (whereas the analysis presented in this chapter extends the model to age 11). As the model requires at least four periods of observation to model the recursive relationship, Hernández-Alava and Popli (2017) used cognitive measures taken in infancy. These early childhood measures were originally designed to identify key developmental milestones and show less variation compared to cognitive measures taken at later stages of development.[88]

Another key difference between the present study and the work of Hernández-Alava and Popli (2017) is the approach used to define parental investment. One of the primary contributions of the model presented in Chapter 3 is that it allows for a more nuanced examination of the role played by different parenting inputs. While the modified model presented by Hernández-Alava and Popli (2017) partially addressed this concern by including several distinct types of parental investment, their analysis provided limited theoretical justification for the types of measures included in the model. Specifically, their approach to defining these inputs relies heavily on EFA and does not explain why they chose to define parental input using both self-reported attitudes towards parenting and specific parenting behaviours.

In response to these shortcomings, I have chosen to apply my modified version of the skill formation model to data from the MCS. This not only provides updated estimates

[88]Further details about the specific measures are provided later in this chapter.

of developmental trajectories in children from the UK, but also allows comparison of the findings from my modified version of the skill formation model to the estimates obtained by Hernández-Alava and Popli (2017). As my analysis uses more recent data from the MCS, I am able to model child development over a longer period and capture these children's progression through the end of primary school. Additionally, my analysis includes a detailed discussion of the identification of the parental input factors in the model; the precise definition of these measures has notable implications for the interpretation of results.

Based on the considerations discussed above, the present study makes two main contributions to the existing literature:

- UK-specific estimates of the dynamic model of skill formation using the most recent data from the Millennium Cohort Study. This nationally representative sample includes girls and non-white children, providing evidence that the model is applicable to different populations.
- Evidence on the effect of family characteristics and of specific parenting behaviours on skill development in children. By specifically measuring parenting behaviours, this approach avoids the pitfalls of assuming that family characteristics such as parental income and education are necessarily a proxy for parental behaviour. Similarly, this model design allows me to separate parental behaviours from parental attitudes; though both may influence child outcomes, it is unrealistic to assume that they capture the same underlying factor as intentions and actions are not always the same.

4.2 Data

As detailed above, this data used in this empirical chapter is taken from the UK Millenium Cohort Study (MCS). Before I explore the specific features of the MCS, I briefly discuss the process of identifying UK data suitable for estimating my theoretical model. Next, I provide a detailed overview of the structure of the MCS and the strategies used in data collection. Within this overview, I outline the specific features of the MCS survey sample and discuss how I have selected the subsample that is used in this chapter. Once this subsample has been defined, I discuss the features of the MCS relevant to my research, including a full analysis of each of the measures included in the skill formation model. I conclude this section by reviewing the demographic characteristics of the analysis subsample and examining the descriptive statistics for each of the measures used in the analysis.

Data Selection

In Section 3.4, I explained why cohort studies are well suited to the empirical applications of my methodological framework.[89] Therefore, to apply this model to a UK context, I began by examining each of the available cohort studies to assess their feasibility. Only a handful of such cohort studies exist: the British Cohort Studies and the Avon Longitudinal Study of Parents and Children. Of these options, the British Cohort Studies are more suitable as they are designed to be nationally representative, allowing them to be used to make conclusions about the entire population. Fortunately, the British Cohort studies contain rich, high-quality data, for relatively large samples. Consequently, I have opted to use the Millennium Cohort Study, which is the most recent study of this kind.

The data contained in the MCS is publicly available and has been used in a variety of educational research papers examining the various predictors of school readiness and skill development in the UK (e.g. Flouri, Midouhas, & Joshi, 2014; Flouri, Midouhas, & Ruddy, 2016; Flouri & Sarmadi, 2016; Girard, Pingault, Doyle, Falissard, & Tremblay, 2016; Hartas, 2011; Kelly et al., 2011). As the MCS is an ongoing study, more data becomes available every few years. This dissertation focuses on the data that was available when this analysis began in 2015.

[89] More information about the definition of a *cohort study* is provided in Chapter 3.

Millennium Cohort Study (MCS)

Survey Description

The Millennium Cohort Study (MCS) is a UK-based longitudinal cohort survey that follows a representative sample of approximately 19,000 children in England, Wales, Scotland and Northern Ireland who were born in 2000–2001. The MCS was designed to create a versatile dataset which would track not only various aspects of these children's development but also contain specific information about their family characteristics, daily life and general wellbeing. To accomplish this, the MCS follows this nationally representative cohort of children throughout their childhood and adolescence, collecting detailed measures at regular intervals. The study was commissioned by the Economic and Social Research Council (ESRC) and has been designed and implemented by researchers at the Centre for Longitudinal Studies (CLS).

The first MCS survey took place in 2001–2002, when the cohort of children were around 9 months old. The sample was revisited and data was collected when the children were three, five, seven, eleven and fourteen years old. Within the MCS literature, each data collection period is referred to as a 'sweep.' Figure 4.1 depicts the general structure and sample for the first five sweeps of the MCS. More detail on the MCS is provided in the MCS User Guide (Hansen, 2014), as well as later in this section.[90]

Figure outlining the survey elements of the MCS removed for copyright reasons.

Copyright holder is Institute of Education.

For original figure, refer to Hansen (2014)

Note: This figure is from Hansen (2014; p. 30) where is appeared as 'MCS Survey Content'

Fig. 4.1 MCS at a Glance

[90] At the time that this analysis was designed, these five sweeps were the only MCS data that had been made publicly available. The data from the sixth sweep of the MCS was released in the summer of 2018. Future work could extend the present analysis to include this additional data.

The first five sweeps of the MCS data contain measures of over a thousand variables. These are varied and include measures of the child's family structure, household income, the child's health and other topics relating to family life and household activities. Although an exhaustive review of all variables included in the MCS would be beyond the scope of this dissertation, I do provide a detailed review of the variables that are included in my analysis, including a discussion of why these measures were chosen.[91]

Data Access

To obtain MCS data, researchers must register with the UK Data Service and submit separate research proposals to specifically request data for each project. Prerequisite for access is agreement to abide by the UK Data Service's *Data Access Policy* (2014).

After receiving approval for my project, I was able to download the data, along with survey documentation from the UK Data Service. Each sweep of the MCS is divided into multiple sections, and each of these sections is stored in a separate data file. To protect the confidentiality of respondents, the MCS identifies children using unique ID numbers. As the ID numbers are consistent across study sections and sweeps, they can be used to link responses across the multiple files. Using STATA, I was able to merge the relevant MCS files and create a longitudinal file to be used for analysis.[92] To comply with CLS security requirements, this data was stored on a password protected computer.

Ethical Considerations

Though I did not directly collect the data, it is important to acknowledge that the present study involved human participants. Below, I discuss the ethical review taken prior to the MCS data collection alongside the ethical considerations I made in my analysis.

Ethics reviews were conducted before each MCS sweep and the sample design was granted ethical approval by various National Health Service (NHS) – Research Ethics Committees (CLS, 2014). At the beginning of the study and before each follow-up survey, the goals and structure of the MCS were reviewed with participants who were also told that any information released to the public would be stripped of identifying information.

As the primary focus of the study is children, who are unable to consent directly, informed consent was obtained from the primary caregiver of the child. Written permission forms were required before each sweep of the survey to ensure that consent was ongoing. Within these forms, specific written consent was sought for researchers to access each family's health, tax and education records and to link this data to the child in question. In addition to the written consent obtained from the caregiver, verbal assent was obtained from the child for any assessment in which they participated. The voluntary nature of the

[91] A full list of the measures contained in the MCS is provided by the CLS and included in Appendix A.1.

[92] Following standard practice, detailed do-files for cleaning and merging the data are not included in this dissertation. These do-files can be provided upon request.

study was essential for ethical research collection; the CLS emphasises that "individuals are able to refuse to participate in any element of a survey or withdraw from the study at any time" (CLS, 2014, p.2).

As the data provided by the UK Data Service is stripped of identifying information, it labels participants using ID numbers. Not only do ID numbers allow for data linkage between the survey sweeps, they also ensure confidentiality for respondents. To protect the identity of individuals in small demographic subgroups, I do not report results for groups that contain fewer than ten respondents. Using secondary data makes it impossible for me to debrief respondents on my specific study, but the results will be provided to the CLS and may therefore be fed back to the participants.

Sample Size and Attrition

The original target sample for the MCS was 20,646 households. The first sweep took place in 2001–2002, when the children were 9 months old and had a response rate of 89.9%, resulting in a sample of 18,552 households. The families who responded in the first sweep were set as the target sample for all future MCS sweeps. This target sample was revisited for follow up data collection when the children were 3 years, 5 years, 7 years, and 11 years old. By the fifth sweep of the MCS — when the children were 11 years old — attrition had reduced the number of respondents to 13,287, representing 69% of the original respondents. This included respondents with incomplete responses to one or more of the sweeps. Table 4.1 outlines the MCS sample size for the first five sweeps.

Table 4.1 Number of Respondents: MCS Sweeps 1–5

Survey Year	MCS Sweep	Child Age (in years)	Target Sample	Number of Respondents	Response Rate
2001/2	1	9 (months)	20,646	18,552	82%
2004/5	2	3	19,941	15,590	78%
2006	3	5	19,244	15,246	79%
2008	4	7	19,244	13,857	72%
2012	5	11	19,244	13,287	69%

Source: This table is a summary of information presented in pages 7–15 of "Millennium Cohort Study: A Guide to the Data Sets" (Hansen, 2014, pp. 7–15).

To increase the sample size of the MCS, the CLS added a 'top-up' sample, of 1,389 families who did not respond to the first sweep, to the second sweep's target sample. These 'new families' were eligible for the original sweep of the MCS but were 'non-respondents' as the available home address was incorrect (Hansen, 2014). Along with the Sweep 1 sample, these families were followed for the remainder of the MCS.[93]

[93] This sample is sometimes mistakenly thought to oversample low-income children. The target sample was unchanged from Sweep 1, but since low-income families have a higher likelihood of changing address, they were more likely be untraceable in Sweep 1; therefore they are overrepresented in the top-up sample.

Sample Design

The MCS used a clustered and stratified sample design, which oversampled electoral wards in Wales, Scotland and Northern Ireland, in areas of financial deprivation, and in areas with high levels of racial diversity. In the selected electoral wards, all children born during a particular period were included in the sample. To account for this sampling strategy, as well as attrition, the sampling weights provided by the MCS are used in all of the analyses in this dissertation. Below, I outline the MCS sampling strategy: complete information can be found in the MCS reference material (Hansen, 2014).

Though the sampling strategy used by the CLS to conduct the MCS was outside of my control, it remains essential to discuss the sample selection before conducting any analysis. The use of a representative sample allows me to obtain estimates that accurately reflect the UK population. As discussed in Chapter 3, representative samples must be designed in such a way that they are not only representative of the larger population of children born in the reference period but also contain large enough subsamples of various subgroups within the population. The size of these subsamples is needed to provide sufficient statistical power in any analysis, which aims to compare the subgroup. To overcome these concerns, the MCS made use of two sampling strategies: stratification and clustering. An explanation of these sampling strategies is provided in Chapter 3, with the MCS specific application of these methods outlined in Appendix A.2.

The use of stratification and clustering is important to consider in any analysis, as the MCS sample is not a simple random sample of the UK population. As a result of this sampling strategy, the raw distribution of characteristics observed in the MCS data may not reflect the actual distribution of characteristics in the UK population. Furthermore, the majority of basic statistical analyses rely on the assumption that a sample is randomly selected, so researchers must be cautious when using non-random samples. Fortunately, there are statistical techniques that can be used to make any findings more representative of the target population and these will be discussed in greater detail within my analysis.

In addition to the non-random design of the MCS sample, the data is subject to attrition and item non-response. In the context of the MCS, item non-response occurs when a child has responses from the parent survey but is missing the survey responses from the child's school. The overall response rate of the MCS is presented in Table 4.1. This includes all individuals who have responded to the main portion of the survey, but does not address those missing other portions. The response rate of 69% in the fifth MCS sweep is considered excellent for a longitudinal study of this size, as response rates are generally lower. Cohen et al. (2007) examine attrition in survey research and find that longitudinal surveys often only yield responses from 20–30% of their original sample.

As stated above, due to the sampling strategy of the MCS, traditional estimates may not reflect the reference population. This problem is magnified by the presence of attrition and item non-response as both of these are unlikely to be randomly occurring. To help researchers address this concern, the MCS data includes final sampling weights that intend to account for the sampling strategy and attrition.[94]

Using the raw administrative data obtained from records from administration of the Child Benefit, the CLS calculated sampling weights for the MCS sample (Plewis, 2014). These weights account for the sample design, but are also re-adjusted for each sweep to account for attrition. Thus, when survey weights are used, the sample characteristics of the weighted MCS sample match known population totals. For example, areas with high ethnic minority populations were oversampled in the original MCS sample. This oversampling ensures that a sufficient number of individuals in these ethnic groups are surveyed. To correct for oversampling, children in these wards would be assigned a lower sampling weight, while children from other wards which were not oversampled are assigned higher sampling weights. In any analysis using sampling weights, estimates are adjusted so that observations assigned higher sample weights are counted more times than those with lower sample weights. This provides point estimates that reflect the true composition of the population; by contrast, a simple point estimate would put too much emphasis on the traits of disadvantaged individuals (as they are overrepresented) and neglect the responses of advantaged individuals (who are underrepresented).

Since the MCS sampling weights are generated using national level characteristics corresponding with Sweep 1, any estimates from an analysis using sampling weights will represent the UK population in 2000–2001. These sampling weights do not account for immigration or emigration that took place since the original sample was selected. Thus, weighted estimates may not capture the true demographics of the corresponding population at the time of the later sweeps. While migration measures can be obtained, it is impossible to specify how this migration would have impacted the MCS sample.

Data Collection

To explain how the MCS fits in my analysis, I begin by describing how the data was collected, and the general types of measures included in the MCS. Later in this chapter, I provide a detailed review of each of the MCS measures used in this dissertation. However, as the data included in the MCS is extensive, the current section focuses on the way the data is collected. Earlier in this chapter, Figure 4.1 provides a summary of the people interviewed for each sweep of the MCS. The majority of the survey questions were posed to the child's primary caregiver — referred to as *'main'* — with additional questions being posed to their spouse or significant other, who is referred to as *'partner.'* Within the MCS, the child is referred to as *'cohort member'* or simply *'child.'*

[94]For further information on the use of sampling weights, the reader should refer to Chapter 3, which provides a detailed explanation as well as a simplified example.

Once the target wards were identified using the sampling strategy outlined above, the parents of all children born in the wards were contacted by post when the child was 7 months old. The letter contained a full explanation of the MCS, and the expectations for those participating in the study. Parents were provided the opportunity to opt out of the study and the letter included information on how to do so. If families did not actively opt out, the MCS visited them in their homes when the child was approximately 9–10 months old to obtain the data included in the first sweep. In each sweep, parents were given a paper and pencil survey, followed by a face-to-face interview. Later sweeps included cognitive and physical assessments of the child, as well as questionnaires distributed to older siblings, teachers and school administrators.

Selection of Relevant Subsample

This analysis makes use of the first five sweeps of the MCS data. This allows use of data which tracks respondents from birth until the end of primary school. Respondents must have responded to each sweep, as the model aims to measure trajectories of growth and cannot be analysed for individuals missing an entire data point.

To obtain a sample that met the requirements of the model presented in Chapter 3, I had to apply several sample restrictions to the full MCS sample. The resulting subsample maximises the amount of available information within the constraints of the existing data. Taking into consideration all of these restrictions, the final usable sample for the present analysis contains 8,355 children. This sample is restricted to include only children who:

- have completed the child assessment portion for the first 5 sweeps (10,298 cases).
- are singleton[95] children (10,034 cases).
- have their primary care-giver present for all interviews up to age 11 (9,951 cases).
- provide information for the relevant demographic measures (9,885 cases).
- respond to at least three of the five behavioural measures in each sweep (8,840).[96]
- have scores for at least 50% of cognitive assessments for each sweep (8,355).[97]

Using a restricted subsample of the larger MCS sample has the potential for introducing bias: this bias occurs if the individuals omitted from the subsample differ in some way from those in the restricted subsample. Table 4.3 compares the restricted subsample to the full sample to ensure there are no notable differences between the two.

[95]Singleton refers to children who are not twins or triplets. There is evidence that multiples have lower scores on cognitive and behavioural tests that cannot be explained by other factors(Ronalds, De Stavola, & Leon, 2005; Rutter & Redshaw, 1991). Excluding multiples from the present analysis avoids this bias.

[96]The estimated factor loadings are based on all five behavioural measures. Estimating an individual's latent factor score using fewer than 50% of these indicators introduces substantial bias.

[97]The number of relevant cognitive assessments varies between sweeps. At least 50% of the indicators are required for valid estimates of the latent construct.

Measures used in Present Analysis

Demographic Measures

Chapter 3 has explored three ways that demographic characteristics can be included in the model: If the characteristic:

1. is thought to reflect a child's initial endowment of skill, it is included in the first stage of the model to control for initial ability.
2. captures either the resources available to the child or a factor known to influence development directly, it is included as a period-specific input to skill development.
3. describes a demographic subgroup for which there are systemic differences in how a specific variable is measured, it is included alongside the biased measure to control for measurement error.

Before introducing the relevant demographic measures contained in the MCS, I outline the specific characteristics that I need to control for in order to capture each of the three pathways through which demographic characteristics influence development. This builds on the general theoretical explanation provided in Chapter 3 with a view to justifying the inclusion of each of these types of covariates in the estimation approach. After outlining the necessary types of measures to include, I then introduce the exact variables from the MCS that are used in the empirical application presented in this chapter. Mathematical details of how these covariates are specified in the present analysis will be provided in Section 4.3

Demographic Characteristics as a Proxy for Initial Ability: While it is impossible to know the innate ability possessed by a child, birth weight can be used as a proxy as it provides insight into genetic endowment as well as into prenatal conditions. Matte, Bresnahan, Begg, and Susser (2001) found that for siblings raised in the same household, birth weight is a strong predictor of childhood cognitive ability.

Demographic Characteristics as Available Resources: There are a variety of family characteristics which correlate with the resources available. Evidence from the literature suggests that maternal education is correlated with better resources available in the home environment, which in turn predicts improved child performance on developmental measures (Dickson, Gregg, & Robinson, 2016). Since I have specifically adapted the model to distinguish between parental resources and the effect of specific behaviours, the analysis includes controls for maternal education. This allows me to identify the effect of certain parental behaviours, over and above the effect of parental skill level. The analysis also controls for single parent households and the number of siblings, as both of these factors influence the time and resources the parent has available per child. Controls are also included for equivalised family income, this captures the differing resources available to families and the known correlation between family income and children's cognitive skills.[98] This differs from the approach of Cunha and Heckman (2008) who include income as one of the measures used to create the latent variable of parental investment, and is therefore one of the major methodological contributions of this dissertation.

Demographic Characteristics and Measurement Error: Finally, though some demographic characteristics do not play a direct role in skill development, there may be systemic differences in the way certain skills are reported across genders and races. For example, in the UK, written test scores of non-English speakers are prone to systematically underestimating these children's true ability, so this empirical application must control for English language ability. There is also some evidence that there are consistent differences in the way identical behaviours are reported across gender and race (Goodman, 1997; Zwirs, Burger, Buitelaar, & Schulpen, 2006). Controlling for these demographic characteristics reduces the measurement error from the differences in how skills are measured across groups.

Table 4.2 lists the measures from the MCS which I include in the analysis to control for the sources of bias discussed above.

Table 4.2 Definitions of Included Covariates

Variable	Description
Time Invariant	
Child's Gender	Binary variable set to equal one if the child is male.
Birth Weight	Child's weight at birth, reported by the main respondent.
Mother's Age at Birth	Determined using mother's reported date of birth.
Language Spoken in Home	Categorical variable for 'English Only', 'English and other language(s)' and 'Other languages only'.
Respondents's Highest Level of Completed Education [1]	Measured when the child was 9 months old. Responses fell into 8 categories.
Child's Ethnicity	Respondent asked to define child's ethnicity. Responses are coded into six categories.
Time Variant	
Single Parent Household	Based on the main survey respondent's description of who lives in the household.
Number of Siblings	Based on the main survey respondent's description of who lives in the household.
Household Income Quintile	Categorical variable, assigning income quintiles based on the family's reported income in the previous calendar year.

Note:
[1] With the exception of 4 children in the sample, the mother was the primary respondent. Thus, this can be interpreted as maternal education.

[98] In the UK context, social class also plays a role. This may not be captured by measures of income.

Though Table 4.2 provides a brief description of the relevant covariates, the text below contains details of exactly how they were measured by the MCS and the way in which they were coded in the present analysis. Demographic variables were measured during all sweeps of the MCS. For those that remain fixed over time the responses from Sweep 1 were used in the analysis. The time-invariant measures are defined as:

- **Child's gender:** as reported on the child's birth certificate. Equal to one for male and zero for female.
- **Birth Weight:** Respondent was asked to consult the child's Personal Child Health Record (commonly known as the "red book") and to report how much the child weighed when she/he was born.
- **Mother's Age at Birth:** This variable was calculated based on the difference between the child's date of birth and the mother's date of birth.
- **Language Spoken in Home**: respondent was asked "can I just check — is English the language usually spoken at home?". Respondent was then requested to indicate whether other languages were spoken at home, being given the choice of 'English only', 'English and other languages' and 'Other languages only'. Two binary indicator variables were created for: 'Other languages only' and 'English and other.'
- **Respondent's Highest Level of Education:** Respondents were asked what was the highest level of education she/he had completed. The interviewer recorded responses using 8 categories. Binary variables were created for each of these categories, equal to one for the highest level completed and zero otherwise. The categories were:
  - *Higher Degree and Postgraduate Qualifications*
  - *First Degree (including B.Ed.)*
  - *Post-graduate Diplomas and Certificates*
  - *A/AS/S Levels*
  - *O Level or GCSE grade A-C*
  - *GCSE or O Level below grade C*
  - *Other academic qualifications (including overseas)*
  - *None of the above*
- **Child's Ethnicity:** the parent was asked how they would describe their child's race. The MCS provides several categorisations based on this response. The present analysis uses the responses coded into six categories. Children were assigned a value of one for the ethnicity in which their parents selected as appropriate for them, and zero for all the other categories. The categories were:
  - *White*
  - *Mixed*
  - *Indian*
  - *Pakistani and Bangladeshi*
  - *Black or Black British*
  - *Other Ethnic Group*

Table 4.3 contains summary statistics of the demographic characteristics for the specified subsample and the full sample from Sweep 1 of the MCS. For the binary variables (ethnicity, primary-respondent education and child's gender), the mean can be interpreted as the percentage of the sample possessing the measured feature.

Table 4.3 Time Fixed Demographic Measures: Unweighted Descriptive Statistics

	Analysis Sample[1]		Full MCS Sample[2]		
	mean	SD	mean	SD	obs
Child's Gender (=1 if male)	0.493	0.500	0.514	0.500	18,552
Birth Weight (kilograms)	3.398	0.567	3.344	0.590	18,487
Mother's Age at Birth	29.301	5.660	28.330	5.966	18,550
English & Other Language Spoken in Home	0.069	0.254	0.112	0.358	18,552
No English Spoken in Home	0.015	0.123	0.039	0.193	18,552
Highest Educational Qualification (Primary Respondent)[3]					
Higher Degree	0.043	0.202	0.033	0.180	18,484
First Degree	0.170	0.376	0.124	0.329	18,484
Post-Grad. Dipl. & Cert.	0.106	0.308	0.084	0.278	18,484
A/AS/A Levels	0.109	0.312	0.093	0.290	18,484
O-Level/GCSE (grades A-C)	0.350	0.477	0.335	0.472	18,484
O-Level/GCSE (grade <C)	0.097	0.296	0.107	0.309	18,484
Other Qual. (Inc. Overseas)	0.016	0.127	0.029	0.167	18,484
None of the Above	0.109	0.311	0.195	0.396	18,484
Child's Ethnicity					
White	0.903	0.296	0.826	0.379	18,504
Mixed	0.024	0.152	0.030	0.170	18,504
Indian	0.019	0.137	0.025	0.157	18,504
Pakistani and Bangladeshi	0.029	0.168	0.068	0.252	18,504
Black or Black British	0.018	0.135	0.036	0.186	18,504
Other	0.006	0.080	0.014	0.119	18,504
Observations	8,355		18,552		

Notes:
[1] Children who: completed the child assessment portion for the first 5 sweeps of the MCS, are singleton children, have a primary care-giver present for all interviews, provide relevant demographic measures and at least 50% of behavioural and cognitive measures.
[2] All responses to the first sweep of the MCS (MCS 1). Some respondents were missing information for some of the measures, so the relevant number of observations is provided for each measure.
[3] With the exception of 4 children in the sample, the mother was the primary respondent. Thus, this can be interpreted as maternal education.

The unweighted sample statistics in Table 4.3 show the true composition of the analysis sample.[99], with slight differences between the analysis sample and the full MCS sample. The analysis sample has higher maternal age at birth, lower levels of non-English speakers and higher levels of education when compared to the MCS sample in Sweep 1. This corresponds with the expected patterns of attrition.

[99] Since sample weights are calculated using demographic measures, weighted descriptive statistics would not reveal the composition of the subsample as the weights are designed to recreate the characteristics of the reference population.

The covariates above remain fixed over a child's lifetime, but the other covariates listed in Table 4.2 change over time and must be measured during each period of the study. The relevant time-variant measures were collected during each MCS sweep and are defined as:

- **Single Parent Household:** The main respondent was asked a series of questions to describe the household structure. These questions were used to define whether 'two parents/carers' or 'one parent/carer' lived in the house. This was used to create a binary variable equal to one if the child lived with 'one parent/carer'.
- **Number of Siblings:** During the questions used to describe the household structure the main respondent was asked about all children living in the household. The total number of siblings included natural, half, step, adopted and foster. This was included as a integer representing the total number of siblings.
- **Family Income Quintile:** The family was asked to report their income in the previous year and this, along with the number of members in the household, was used to calculate the equivalised household income. The equivalised income was based on modified OECD[100] scales for equivalisation, which sets a couple with no children as equal to one and assigns ratios to families of other sizes. Further information on the income equivalisation calculations are provided in Hansen (2014). This equivalisation accounts for both income and family size, and was used to place the family into one of five income quintiles based on the entire MCS sample.

  The decision was made to use the MCS reported quintiles instead of recalculating using the specific subsample. Using the original quintiles maintains a clearer definition of a family's position in the income distribution because the subsample may not be as balanced a representation of the true income distribution. In the analysis, children were assigned a score of one for the quintile to which they belonged, and zero for all the other quintiles.

Table 4.4 provides the unweighted descriptive statistics of the analysis sample for the time varying covariates relevant to this analysis. As with Table 4.3, the use of unweighted descriptive statistics in Table 4.4 allows the reader to see how these covariates are distributed in the analysis sample. Unlike Table 4.3, it is not possible to compare the analysis sample to the full MCS sample for each of these measures, because not all respondents in the full sample completed the later MCS sweeps.

As expected, Table 4.4 shows that the number of siblings rises as the child ages, indicative of the cohort children's new siblings born as time progresses. Similarly, the rate of single parent families increases slightly as the child ages, a pattern that is in line with

[100] The Organisation for Economic Co-operation and Development (OECD) is an international economic association. The OECD income scales provide standardised methodologies which allow for household income to be adjusted for the number of individuals in a household.

Table 4.4 Time Varying Demographic Measures: Unweighted Descriptive Statistics

	MCS 1 *Age 1*		MCS 2 *Age 3*		MCS 3 *Age 5*		MCS 4 *Age 7*	
	mean	SD	mean	SD	mean	SD	mean	SD
Single Parent	0.124	0.329	0.139	0.346	0.160	0.367	0.178	0.383
No. of Siblings	0.872	0.977	1.144	0.992	1.332	0.983	1.461	1.001
Household Income[1]								
Lowest Quintile	0.162	0.368	0.161	0.368	0.154	0.361	0.151	0.358
Second Quintile	0.192	0.394	0.190	0.392	0.191	0.393	0.183	0.387
Third Quintile	0.209	0.407	0.213	0.409	0.215	0.411	0.215	0.411
Fourth Quintile	0.223	0.416	0.221	0.415	0.222	0.415	0.226	0.418
Highest Quintile	0.214	0.410	0.215	0.411	0.218	0.413	0.225	0.418
Observations	8,355		8,355		8,355		8,355	

Note:
[1] Income quintiles are determined using weekly household income adjusted for family size using the OECD equivalence tables. For more information on their derivation, see MCS user guide.

general demographic trends. The descriptive statistics show a slight underrepresentation of those in the lowest income quintile. This is likely to result, at least in part, from the decision to include only respondents who had completed all five sweeps of the MCS; those in the lowest quintile are more likely to have missed part of the survey (Plewis, 2014). The use of the MCS-provided sample weights will help control for this sample bias in the analysis conducted for this dissertation.

Measures of Cognitive Ability

At ages 3, 5, 7 and 11, the MCS conducted a variety of cognitive assessments that were administered directly to the child by the interviewer. Table 4.5 lists the available cognitive measures from all five sweeps of the MCS. Although it would be ideal to have the same measures repeated over time, this is not possible within the context of the MCS. Fortunately, extensive literature exists which discusses how the various BAS measures can be compared to each other, along with comparisons to the Bracken School Readiness assessments and how these scores can track a child's development over time.

Table 4.5 Cognitive Measures Included in the MCS

	MCS Sweep			
	MCS 2 *Age 3*	MCS 3 *Age 5*	MCS 4 *Age 7*	MCS 5 *Age 11*
BAS Naming Vocabulary	X	X		
Bracken School Readiness	X			
BAS Picture Similarity		X		
BAS Pattern Construction		X	X	
BAS Word Reading			X	
BAS Verbal Similarities				X
NFER Number Skills			X	
CANTAB Spatial Working Memory Task				X
CANTAB Cambridge Gambling Task				X

Notes: BAS—British Abilities Scale; NFER—National Foundation for Educational Research; CANTAB—Cambridge Neuropsychological Test Automated Battery
Source: 'MCS Cognitive Assessments by Sweep Collected' (Hansen, 2014; p.62)

For the latent factor of cognitive ability, I use the British Ability Scale scores, the Bracken School Readiness assessment and the National Foundation for Education Research — Progress in Maths test. These correspond with $Y^C_{j,t}$ in Equation 3.11 and are described below. I chose to omit the CANTAB tasks as these two tests have a smaller evidence base in the existing literature, and the MCS data guides caution against using these tests for research that is to be generalised beyond the MCS (Hansen, 2014).

Bracken School Readiness Assessment (BSRA): The BSRA was administered when the MCS children were 3 years old. This assessment measures 88 concepts and involves an oral assessment of the child by the MCS interviewer. The BSRA is an adaptation of the more widely used Bracken Basic Concept Scale—Revised (BBCS–R). The BBCS–R is designed to assess the development of basic cognitive concepts in children ages 2-years-6-months to 7-years-11-months.

Though it was designed to measure school readiness, the BBCS is highly correlated with other measures of cognitive ability. When comparing the BBCS with the Wechsler Preschool and Primary Scale of Intelligence—Revised (WPPSI–R), Laughlin (1995) found a correlation of 0.77 between the two scales. This is similar in magnitude to the correlation amongst scores on major comprehensive intelligence tests (Carvajal et al., 1991). While the BBCS and traditional intelligence measures are highly correlated, the BBCS is much better suited to large surveys as it can be administered in under ten minutes (Bracken, 2002; Hansen, 2014). This is a sharp contrast to major comprehensive measures of childhood cognitive ability, which require a trained professional and over an hour to administer. For this reason, the BBCS is widely used to assess school readiness.[101]

The BBCS may not equally predict the ability of all children. Studies by Panter (2000) and Panter and Bracken (2009) evaluate the validity of the BBCS for predicting success in American children during their first year of school and find that, while the test is a good predictor of outcomes, it performs differently depending on the child's racial/ethnic status. This indicates that any analysis using the BBCS must include demographic controls to account for systematic measurement error.

The MCS adaptation of the BSRA focuses on the first six areas of the BBCS–R, which Bracken (2004) defines as the School Readiness Composite. These six areas are: colours, letters, numbers/counting, sizes, comparisons and shapes. The number of correct answers on these six subtests are combined to provide a composite score for the child, which is reported as a percentage mastery.

The MCS also reports an age-adjusted, standardised score derived using norming tables in the BSRA manual. These norming tables are devised using a representative sample of the US population to set a mean score of 100 and standard deviation of 15 for the test (Bracken, 2004). To avoid bias from using a norming sample from the US, the present analysis uses the percentage mastery scores and adjusts for age within the model.

British Abilities Scales (BAS): A revised version of BAS, the British Ability Scales—2nd Edition (BAS–II) is administered verbally to the MCS children at ages 3, 5, 7 and 11. This version was introduced in 1996 by Elliott, Smith, and McCulloch (1996). Though the MCS uses the BAS–II, for the remainder of the dissertation it is referred to as the BAS. Included in the MCS are five BAS measures. These subtests are pattern construction, naming vocabulary, picture similarity, word reading, and verbal similarity.

The British Ability Scales were introduced in 1979 to serve as standardised assessments of cognitive ability and educational achievement.[102]

[101] Bracken and Panter (2011) provides a review of the BBCS's historical applications and general validity. Further information on the BSRA in the MCS and the BBCS–R in general can be found in Hansen (2014) and Bracken (2002, 2004) respectively.

[102] The BAS was originally normed using a sample of British school children, but the test was later re-scaled using an American standardisation and released as the Differential Ability Scales or DAS (Sparrow & Davis, 2000). Although the sample used to scale the DAS differs from that of the BAS, the testing materials were very similar, and the creators of the BAS were involved in the process of re-releasing the test in its American format. As a result of the two versions of the test, much of the discussion of the validity of the BAS refers to the DAS, but this is merely semantic and Elliott et al. (1996) use DAS and BAS interchangeably in the documentation accompanying the BAS–II.

The BAS are ideally suited to my research agenda, as they have been designed to track cognitive ability over time. Furthermore, there is extensive literature showing that each sub-score has sufficient specificity to be interpreted individually (see Dunham, Mcintosh, & Gridley, 2002). This allows a comparison between the multiple subtests that might not be possible if a measure of general cognitive ability were to be used.

An added benefit of the BAS is that the original test was standardised using a representative sample of British children, which is unusual as the majority of tests have been normed with American samples (Hill, 2005). This standardisation provides an additional method of testing the validity of the assumption that the MCS is a representative sample of the British population.[103]

Included in the MCS are five separate BAS measures, several of which are repeated across sweeps. These subtests are pattern construction, naming vocabulary, picture similarity, word reading, and verbal similarity. The pattern construction task requires children to replicate a pattern using blocks with different coloured sides and assesses spatial awareness. In the naming vocabulary task, which tests expressive verbal ability, the child must name objects pictured in a booklet. The picture similarity task is designed to measure problem solving ability and has the child complete a set of four pictures by choosing a fifth picture that is most similar to others. The word reading task has a child read aloud a list of words and assesses English reading ability. The verbal similarity task asks children to identify the similarity between three words and assesses verbal skills and vocabulary. Further specifics on the BAS tests can be found in Elliott et al. (1996).

Each of the BAS subtests are scored differently. The scores are based on correct answers, speed and understanding. Once the raw scores are obtained, the look-up tables provided in the BAS–II Scoring Manual (Elliott et al., 1996) can be used to obtain scaled ability scores. These adjust for the difficulty of questions asked, and time taken to respond. Finally, norming tables are also used to compare the child's ability to a mean score of children of the same age in the standardisation sample. This results in standardised scores which are reported using a mean of 50 and standard deviation of 10 — with the exception of word reading, which has a normed score of 100 and a standard deviation of 15. As such, for all other BAS scales a child with a score of 60 would be one standard deviation above the mean of their age group.

NFER Progress in Maths (PiM): At age 7 the MCS assessed children using an adaptation of the standard Progress in Maths (PiM) test from the National Foundation for Education Research (NFER). The PiM test was designed to assess numbers, shapes, measurement and data handling. To shorten the time required for assessment, the MCS had the children complete an initial assessment, which was used to assign an 'easier',

[103] For example, if the children in the MCS subsample score poorly on the BAS it might indicate that the sample is of lower cognitive ability than the UK reference population used by Elliott et al. (1996).

'medium' or 'harder' secondary subtest.

PiM scores are reported as scaled raw scores, which are calculated based on the difficulty of test administered and number of correct answers. The MCS also provides age-standardised scores, with a mean of 100 and standard deviation of 15. These standardised scores are created using norming tables based on a nationally representative UK sample.

Unlike the BSRA and the BAS which were specifically designed to be a more general measure of ability, the PiM test is specifically focused on mathematics knowledge. For this reason, caution should be exercised before using this test as a sole measure of ability as it neglects verbal aspects of intelligence. Fortunately, the PiM is administered in the same sweep as the BAS, so the two together will give a better picture of overall ability.

Cognitive Scores in the MCS Sample

Table 4.6 provides summary statistics for both age-standardised and non-age-standardised scores for the cognitive measures. The present analysis uses scores that have not been adjusted for age, and controls for age within the measurement model. This avoids any bias from the norming samples and allows for a consistent age-standardisation across all three types of test. The age-standardised scores are presented below to show how the MCS sample compares to the norming samples used to create the tests.

Table 4.6 Measures of Cognitive Ability: Unweighted Descriptive Statistics

	Age-Standardised Scores		Scores without Age Standardisation		
	Mean	SD	Mean	SD	Obs
Age 3 — MCS Sweep 2					
Bracken School Readiness	105.976	15.683	30.145	15.298	7,878
BAS Naming Vocabulary	51.143	10.769	75.822	16.421	8,243
Age 5 — MCS Sweep 3					
BAS Naming Vocabulary	55.968	10.282	110.387	14.489	8,353
BAS Picture Similarity	56.428	10.000	83.135	11.094	8,354
BAS Pattern Construction	51.596	9.537	89.685	17.931	8,336
Age 7 — MCS Sweep 4					
NFER Number Skills (PiM)	99.405	15.194	19.034	5.514	8,339
BAS Word Reading	112.910	17.606	108.914	29.805	8,217
BAS Pattern Construction	54.298	10.563	118.109	15.688	8,317
Age 11 — MCS Sweep 5					
BAS Verbal Similarities	59.642	9.467	122.185	15.776	8,355

Note: Standardised scores reported in this table are age-standardised based on the reference populations for the respective tests.

For the BSRA, the MCS population has a sample mean of 105.976 which falls above the reference population's mean of 100. This might indicate that the American norming population differs slightly from the UK population used to create the MCS sample. Alternatively, it may mean that the MCS sample is more advantaged than both the US norming sample and the general UK population. Similarly, the sample means for all the BAS measures are above 50, which indicates the sample outperforms the norming sample used by the BAS. Most concerning is the BAS word reading score which has a mean of 112.910, which would indicate the MCS sample mean is nearly an entire standard deviation above the norming sample used for the test. Upon closer inspection the distribution of these scores shows a significant right censoring effect of the age-standardised scores, which results from a large portion of the children falling on the lower end of the age-range for the test administration. Finally, the mean for the NFER falls slightly below the standardised mean of 100, which indicates that, on this measure, the sample performs slightly below the norming population.

Histograms of the relevant variables are shown in Figure 4.2, Figure 4.3, Figure 4.4 and Figure 4.5. The left column of each figure shows the age-standardised scores, while the right column shows the scores that have been scaled for difficulty but not for age.

Unsurprisingly, the age-standardised scores follow a normal distribution. This reflects the fact that the norming tables are intended to create a normal distribution. Many of the scale scores are also normally distributed, indicating that the difficulty scaling also yields a score distribution with higher concentrations near the sample mean.

These histograms highlight that many of the age-standardised scores have either a floor or ceiling effect, demonstrated by a clustering of scores near the bottom or top end respectively. The word reading scores in Figure 4.4 provide the clearest example of a ceiling effect on the standardised test scores. This is captured by the clustering of children with a score of 145. If the censoring was seen in the non-standardised scores, such clustering would indicate that the test is censoring children whose ability is outside the range of the test. Similarly, at the bottom end, such censoring might indicate that these children did not understand the test, or that they were unable to answer any questions, while censoring at the top end indicates that the test may be too easy for the child. However, as the censoring predominantly appears in the age-standardised scores, it indicates that in order to standardise across ages, top scores are awarded to a range of raw scores within this age range, therefore losing some of the information contained in the raw scores. Both of these types of censoring create potential problems with any modelling as they fail to differentiate the ability of the children at the extremes.

To avoid the bias created by the age-standardised scores resulting in a cutoff point for high- and low-ability children, this dissertation uses the scores without age-standardisation and controls for age separately in the model. This approach allows me to benefit from the full distribution of the test scores while still controlling for age.

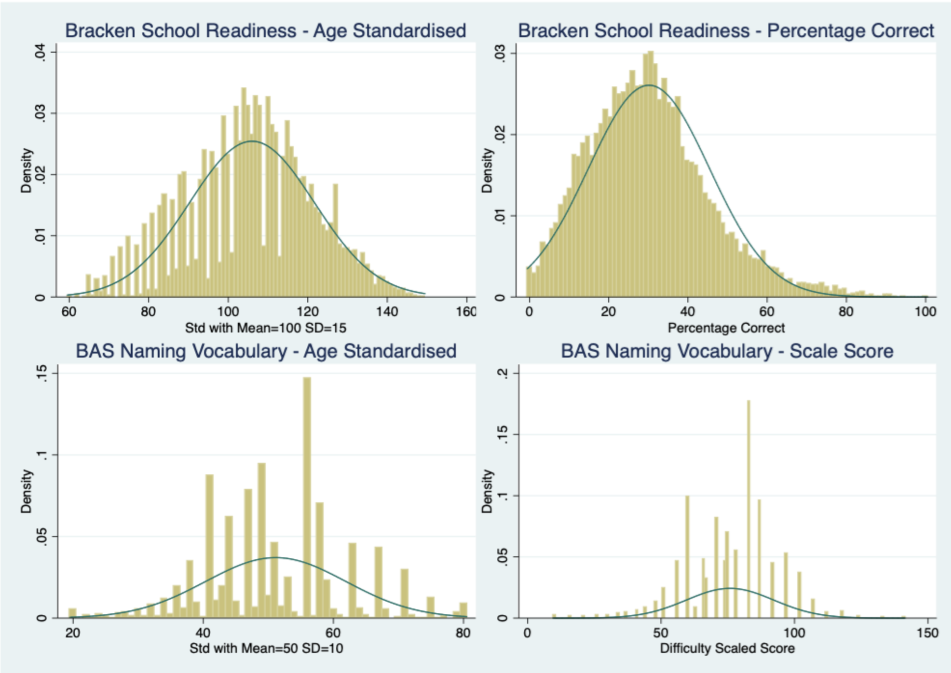


Fig. 4.2 Distribution of Cognitive Scores: MCS Sweep 2 (Age 3)

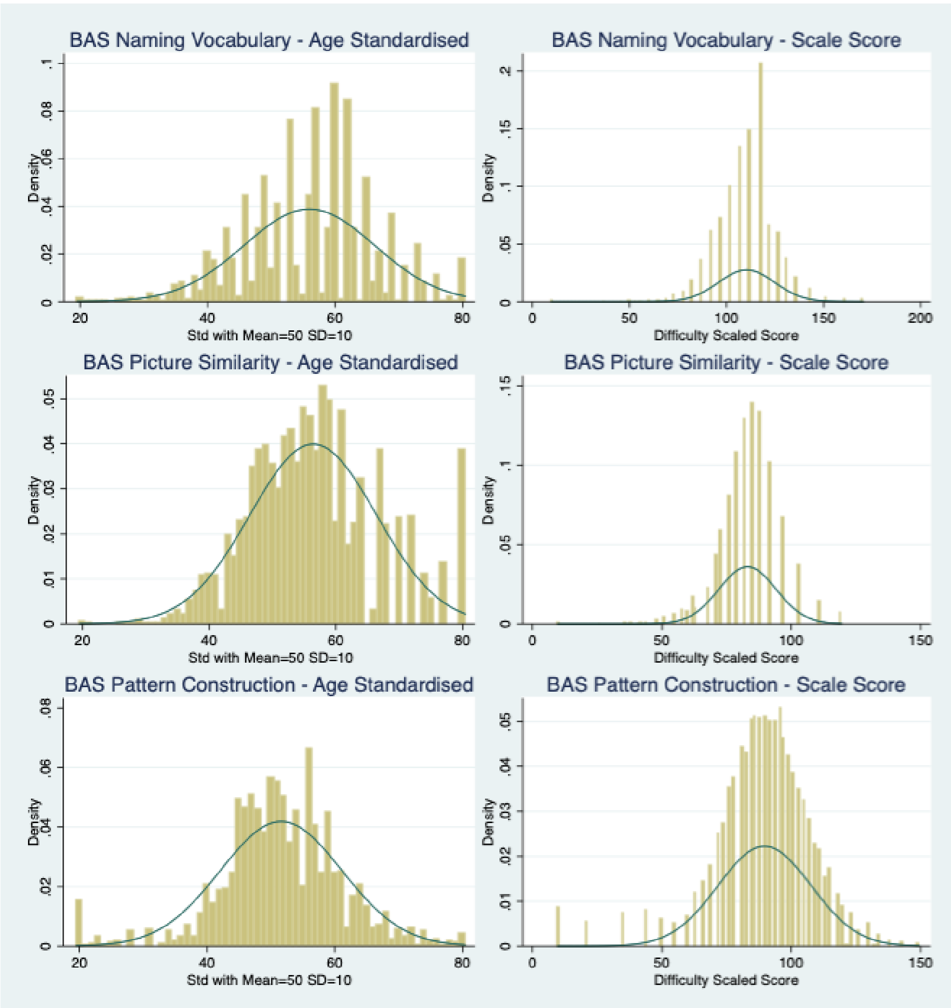


Fig. 4.3 Distribution of Cognitive Scores: MCS Sweep 3 (Age 5)

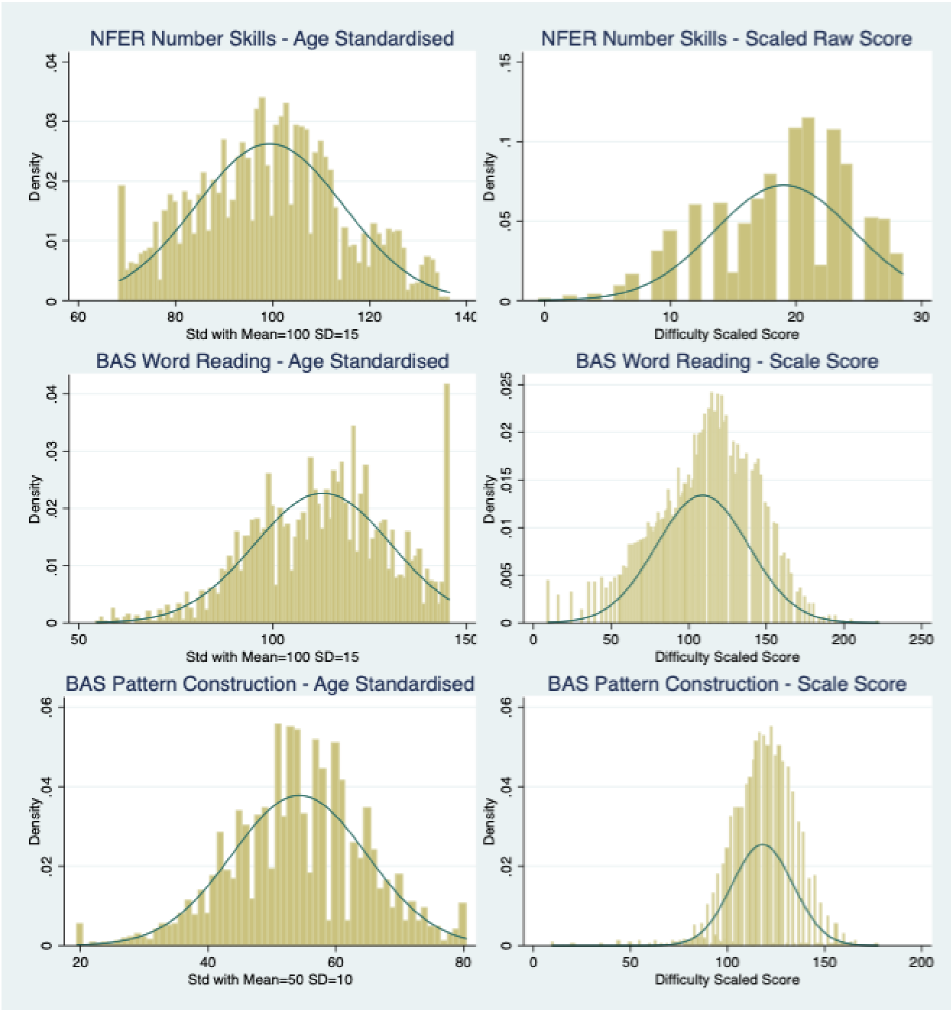


Fig. 4.4 Distribution of Cognitive Scores: MCS Sweep 4 (Age 7)

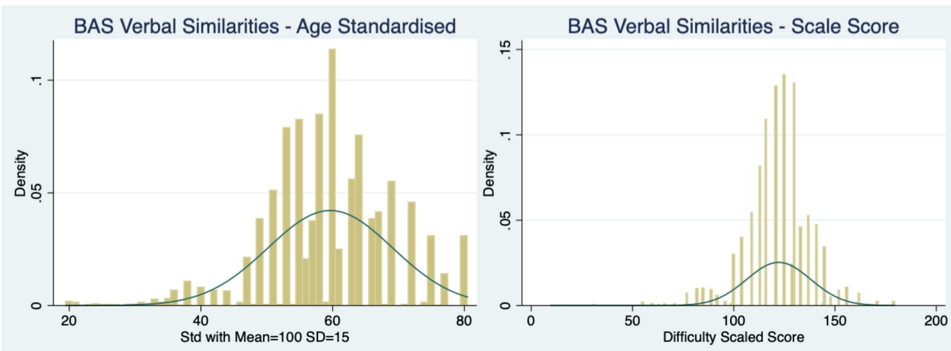


Fig. 4.5 Distribution of Cognitive Scores: MCS Sweep 5 (Age 11)

Measures of Non-Cognitive Ability

As discussed in earlier portions of this dissertation, non-cognitive abilities are not as clearly and consistently defined in the literature as cognitive abilities. Section 2.1 has explained why I have chosen to use a definition of non-cognitive ability which describes measures of behaviour, social skills and emotional development that capture personality traits and emotional intelligence.[104,105]

For non-cognitive measures, I am fortunate that the MCS includes several measures of behaviour, social and emotional abilities across the multiple sweeps. The MCS itself includes parent-reported behaviours as well as providing the option to link English respondents to the data in the Foundation Stage Profiles (FSP). The FSP are a set of mandatory assessments that are conducted in schools in order to assess how a child is doing in formal education. Included with the MCS are measures of a child's achievement as reported by their teacher at the end of the first year of school. Generally, children are 5 years old at the time of this assessment. Although the measures of the FSP are very detailed, the linked data lacks the sufficient sample size to meet my research needs. Furthermore, using the FSP would restrict my analysis to English respondents only. This sample restriction is undesirable, especially as other suitable measures are present.

For the purpose of my research, I have chosen to focus on the behavioural measures contained in the Strengths and Difficulties Questionnaire (SDQ) as this measure is repeated in each sweep of the MCS. Not only does the SDQ allow me to see how certain behaviours evolve over time, but using this measure allows me to use a larger sample size than would be available were I relying solely on the FSP or other measures. Furthermore, the SDQ has been used in multiple studies beyond the MCS; this helps confirm its construct validity and allows me to compare my sample to other populations.

The SDQ was introduced by Goodman (1997) in order to provide an easily administered screening tool which could identify children at risk of psychiatric pathology. This screening tool asks an informant how well each item, contained in a list of 25 attributes, describes the child. These attributes can be divided into five categories of behaviours: hyperactivity-inattention, peer problems, pro-social behaviour, conduct problems and emotional symptoms. For each category, a score out of ten points is given. These scores can be combined to provide an overall measure of a child's psychiatric functioning, or each section can also be individually interpreted. Each of these categories is defined as a separate *SDQ scale* and is based on five attributes from the questionnaire. For

[104] It is worth repeating that this definition of non-cognitive ability does differ from the way this term is used in psychology literature, but that this definitional variation would not change the application of my empirical model.

[105] This matches the way that non-cognitive ability is defined in the skill formation originally presented by Cunha and Heckman (2007). As presented in Chapter 3, my empirical chapters intend to apply an adapted version of their model to obtain a novel set of estimates. Having corresponding definitions is a key part of this approach.

each SDQ scale, children who fall above a prescribed 'cutoff' are thought to warrant further evaluation on that scale. These subscales were originally based on theoretical constructs, but Goodman (2001) used EFA to confirm that the five-factor structure held in a nationally representative sample of British 5–15 year olds.

Though originally conceived as a screening device for pathological behaviours, there is sufficient evidence that the SDQ is still a reliable measure of the distribution of certain traits in the non-clinical population (Meltzer, Gatward, Goodman, & Ford, 2000; Stone et al., 2015). Goodman, Ford, Simmons, Gatward, and Meltzer (2000) note that there is sizeable variability of scores at the subclinical level and that this is indicative of variation in development. Children may demonstrate several behaviours without being above the clinical cutoff, and the scores on the SDQ have significant variation to indicate differing levels of problematic behaviours. Furthermore, there is substantial evidence showing that the SDQ compares robustly to other measures of behaviour in children. For example, a comparative study by Goodman (1997) finds a strong correlation between SDQ scores and scores on the Rutter Total Deviance Scales with correlation coefficients of r = 0.88 and r = 0.92 for parents and teachers respectively. A meta-analysis of 15 studies by Stone, Otten, Engels, Vermulst, and Janssens (2010) confirms the concurrent validity of the SDQ and argues that: "the five-factor structure was confirmed by [15 out of 18] studies, correlations with other measures of child psychopathology were high, and evidence for the screening ability of the SDQ was convincing" (p. 268.)

As well as the strong construct validity of the SDQ, Stone et al. (2010) note that the SDQ's brevity and simplicity of administration have made it a popular choice when compared to other screening tests. This screening tool can be administered at various stages in a child's development and is based on the observations of parents or other caregivers. Duncombe et al. (2012) argue that a key strength of the SDQ is the ability to use the same questionnaire for various adults in a child's life (e.g. parents, teachers, caregivers). The extensive use of the SDQ has resulted in a large body of research discussing the measure's validity. The meta-analysis of Stone et al. (2010) finds that the SDQ "ratings showed sufficient reliability over time, and agreement between parents and teachers was relatively high" (p. 268).

For the MCS, the SDQ is included in the paper-and-pencil questionnaire that is given to the main respondent during each sweep of the survey. The respondent — usually the primary caregiver — was asked to decide if the 25 attributes could be considered "not true", "mostly true" or "certainly true" when used to describe their child.

The statements as included in the MCS questionnaire are provided in Table 4.7. Though presented in order of their subscales, during the survey administration the questions were not asked in this specific order.

Table 4.7 Non-Cognitive Measures: Strengths and Difficulties Questionnaire

SDQ Subscale	Observed Behaviours (e.g. Child is . . .)
Hyperactivity	– restless, overactive, cannot stay still for long – constantly fidgeting – easily distracted – can stop and think before acting* – sees tasks through to the end*
Peer Problems	– tends to play alone – has at least one good friend* – generally liked by other children* – picked on or bullied by other children – gets on better with adults.
Pro-Social Behaviour	– considerate of others' feelings – shares readily with others – helpful if someone is hurt, upset or ill – kind to younger children – often volunteers to help others.
Conduct Problems	– often has temper tantrums – generally obedient* – fights with or bullies other children – can be spiteful to others – often argumentative with adults
Emotional Symptoms	– complains of headaches/stomach-aches/sickness – often seems worried – often unhappy – nervous or clingy in new situations. – has many fears, is easily scared.

Note: The statements marks with * are scored using a reverse of the usual scale.
Source: Table adapted from MCS Guide to the Data Sets (Hansen 2014; p 77)

The MCS datasets use a scale from 0 to 10 to report the score for each SDQ scale. This score is calculated by awarding 0, 1 or 2 points for each of "not true", "mostly true" or "certainly true", respectively. Higher scores correspond with a higher incidence of the behaviours relating to hyperactivity, conduct problems, peer problems and emotional symptoms, and higher scores corresponding with a lower incidence of pro-social behaviours. Though the MCS datasets include the responses for each individual question, I use the subscale scores in my research. For the analysis, controls have been included to capture the respondent's age at the time of interview. Using the scores in their subscale form allows me to compare the results of my study to other research using the SDQ.

The summary statistics for the SDQ are presented in Table 4.8 while Figure 4.6, Figure 4.7, Figure 4.8, and Figure 4.9 provide histograms showing the distribution of the SDQ scores. The summary statistics allow for a sense of how the variables evolve over time, while the histograms allow a more detailed examination of the distributions.

Table 4.8 Measures of Non-Cognitive Ability: Unweighted Descriptive Statistics

	MCS 2 *Age 3*			MCS 3 *Age 5*			MCS 4 *Age 7*			MCS 5 *Age 11*		
	mean	SD	obs.	mean	SD	obs.	mean	SD	obs.	mean	SD	obs.
SDQ Scales[1]												
Hyperactivity[2]	3.722	2.289	8,281	3.102	2.303	8,326	3.178	2.449	8,340	2.959	2.403	8,349
Peer Problems[2]	1.441	1.541	8,295	1.036	1.361	8,330	1.094	1.466	8,340	1.237	1.611	8,353
Emotional Symptoms[2]	1.271	1.414	8,333	1.282	1.513	8,346	1.434	1.686	8,340	1.780	1.935	8,350
Conduct Problems[2]	2.689	1.997	8,345	1.391	1.419	8,353	1.266	1.450	8,352	1.291	1.501	8,353
Pro-Social Behaviour[3]	2.626	1.840	8,297	1.547	1.599	8,351	1.309	1.544	8,353	1.118	1.457	8,354

Notes:

[1] Reported as sum of how well five statements describe the child, each statement receiving a score of 0, 1 or 2.
[2] Higher scores indicate a higher incidence of problematic behaviours.
[3] Scale is reversed for pro-social behaviour with higher scores indicating fewer positive behaviours.

Table 4.8 demonstrates that hyperactivity scores show a slight decrease over time, with the standard deviation remaining relatively constant. Peer problems appear to fluctuate over time, with no clearly discernible trend. There is a slight increase in emotional symptoms as the sample ages, while conduct problems and pro-social behaviour scores show a significant drop between ages 3 and 5 before stabilising from ages 7 to 11.

Figure 4.6, Figure 4.7, Figure 4.8, and Figure 4.9 show that none of the scores are normally distributed and that the distribution of conduct problems and pro-social behaviours appears to narrow as the children get older. This non-normal distribution is to be expected and is indicative of the design of the SDQ as a screening device. As the SDQ is designed to identify problematic traits, we would not expect the frequency of these behaviours to follow a normal distribution. Similar patterns of responses in the British population are discussed in Goodman (2001). The implications of non-normality on the present study are discussed in greater detail in the results section of this chapter.

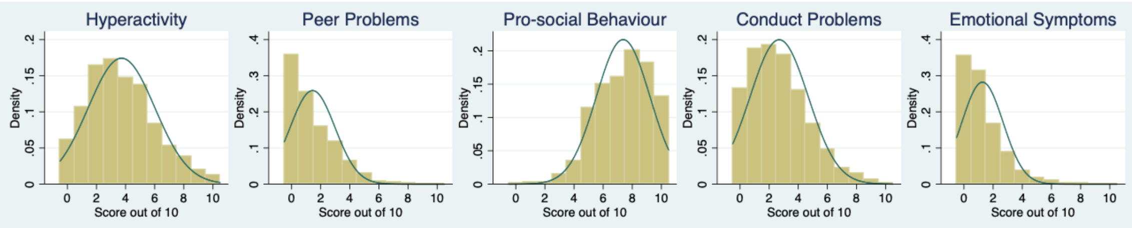


Fig. 4.6 Distribution of SDQ Scores: MCS Sweep 2 (Age 3)

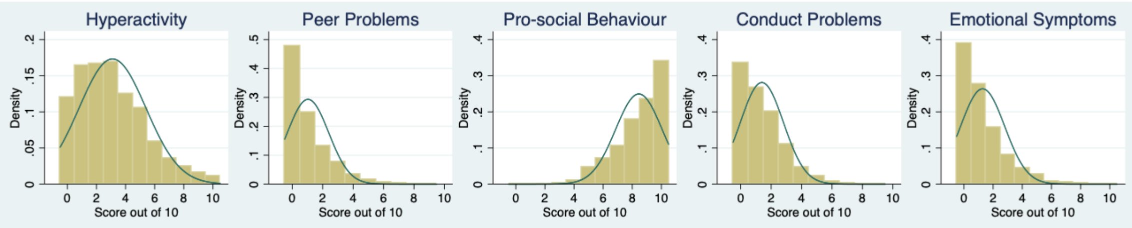


Fig. 4.7 Distribution of SDQ Scores: MCS Sweep 3 (Age 5)

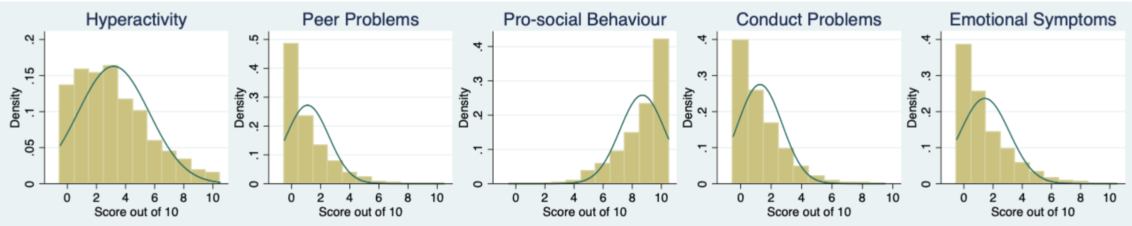


Fig. 4.8 Distribution of SDQ Scores: MCS Sweep 4 (Age 7)

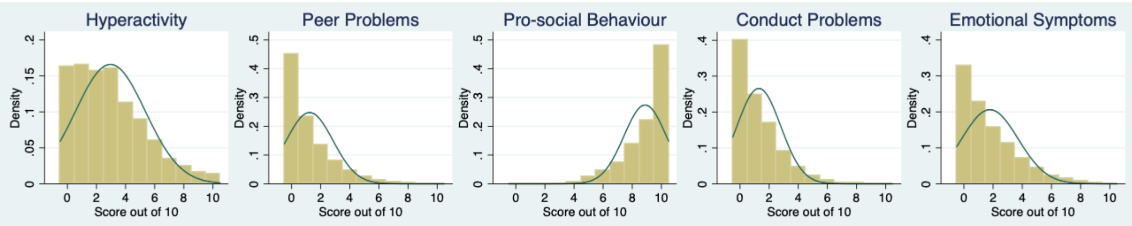


Fig. 4.9 Distribution of SDQ Scores: MCS Sweep 5 (Age 11)

Measures of Parental Investment

To identify the parenting indicators, I relied on a data-driven approach. Upon examining the MCS data, I decided that the construct for which the data contained sufficient indicators is ‘parenting behaviours that require time spent directly with the child’. Using this construct, I identified 13 relevant measures in the MCS, each capturing the frequency that the mother spends engaging in various activities with her child at ages 3, 5 and 7. These activities are listed in Table 4.9, alongside the exact wording of the question used for each sweep.

Table 4.9 Parental Investment Measures: MCS Parenting Questions

‘How often ...’	
1.	... does someone at home take [the child] to the library? *[MCS 2]* *
	... has [the child] been to a library (not a school library)? *[MCS 3,4]* †
2.	... do you read to [the child]? *[MCS 2,3]* ‡
	... do you read *with* or to [the child]? *[MCS 4]* ‡
3.	... does [the child] draw or paint at home? *[MCS 2]* §
	... do you draw, paint or make things with [the child]? *[MCS 3,4]* ‡
4.	... does someone at home try teach [the child] songs, poems or nursery rhymes?*[MCS 2]* §
	... do you play music, listen to music, sing songs or nursery rhymes, dance or do other musical activities with [the child]? *[MCS 3,4]* ‡
5.	... does someone at home help [the child] learn the ABC or the alphabet? ‡
6.	... does someone at home try to teach [the child] numbers or counting? ‡
7.	... do you play sports or physically active games outdoors or indoors with [the child]? ‡
8.	... do you play with toys or games indoors with [the child]? ‡
9.	... do you take [the child] to the park or to an outdoor playground? ‡
10.	... do you tell stories to [the child] not from a book? ‡
11.	... does anyone at home help [the child] with reading? ‡
12.	... does anyone at home help [the child] with writing? *[MCS 3]*
	... does anyone at home help [the child] with writing *or spelling*? *[MCS 4]* ‡
13.	... does anyone at home help [the child] with numbers, counting or adding up? *[MCS 3]* ‡
	... does anyone at home help [the child] with maths? *[MCS 4]* ‡

Notes: Reporting scales are listed below and correspond with the questions as marked above:
* 0 “Less Often or Never”, 1 “On Special Occasions”, 2 “Once a Month”, 3 “Once a Fornight”, 4 “Once a week”
† 0 “Less Often or Never”, 1 “At Least Once a Year”, 2 “Every Few Months”, 3 “At Least Once a Month”, 4 “Once or Twice a Week”, 5 “Several Times a Week”, 6 “Every Day or Almost Every Day”
‡ 0 “Not at All”, 1 “Less than once a month”, 2 “Once or Twice a Month”, 3 “Once or Twice a Week”, 4 “Several Times a Week”, 5 “Every Day”
§ 0 “Never”, 1 “Occasionally or Less Than Once a Week”, 2 “Once or Twice a Week”, 3 “Three Times a Week”, 4 “4 Times a Week”, 5 “5 times a week”, 6 “6 Times a Week”, 7 “7 Times a Week or Constantly”

In Table 4.9, it can be seen that several of the questions changed slightly as the child aged in order to reflect the more age-appropriate behaviour. For example, at age 5, the parents were asked about helping with numbers and counting, while at age 7 this was changed to helping with maths to reflect the change in the type of work a child would be receiving from school. Similarly, the range of possible scores for the scales varied for some of the measures. In each sweep, parents were asked to select, from a list of options, the frequency of the behaviour in question, with the number of response categories varying for different questions. In the analysis, the questions are used in their original frequency scales in order to capture the distribution of their varying scales. [106]

In addition to the 13 measures identified in Table 4.9, the MCS contains other measures that relate to parental investment. These include various features of the home environment, and activities that the child is involved in. While it would have been possible to include measures such as "books in the home", "access to play equipment", or "visits the theatre", such measures do not necessarily indicate a child's time spent with their parent, but rather shows that certain items are present, or that someone takes the child to certain cultural activities. Moreover, such measures may simply be indicators of household wealth or family income, which should be separately controlled for.

Summary statistics for the relevant parental input measures are presented in Table 4.10. With the exception of the frequency of library visits, the summary statistics for all other variables are presented on a 0–5 scale, with a score of 5 indicating the activity occurs daily. These variations show that over the course of a child's first 7 years, parents alter their investment decisions. The standard deviations are not large, but they demonstrate that there is variation between the parenting behaviours of different families.

Table 4.10 Parental Investment Measures: Unweighted Descriptive Statistics

	MCS 2 *Age 3*			MCS 3 *Age 5*			MCS 4 *Age 7*		
	mean	SD	obs.	mean	SD	obs.	mean	SD	obs.
Child visits the Library [1]	1.029	1.314	8,355	1.134	0.980	8,355	1.118	0.955	8,354
Reads to Child [2]	4.346	1.035	8,355	4.288	0.949	8,355	3.990	1.166	8,355
Draws/Paints with Child[2]	4.175	0.967	8,355	2.880	1.169	8,355	2.299	1.198	8,355
Songs/Poems/Rhymes with Child[2]	4.277	1.072	8,355	3.831	1.205	8,354	3.474	1.446	8,355
Helps Child Learn Alphabet[2]	2.952	1.694	8,355						
Teaches Child Counting[2]	4.158	1.128	8,355						
Physical Activities with Child[2]				2.610	1.280	8,354	2.324	1.341	8,352
Indoor Games/Toys with Child[2]				3.544	1.125	8,353	2.915	1.185	8,354
Take Child to Park/Playground[2]				2.668	1.011	8,353	2.430	1.104	8,354
Tell Stories to Child[2]				2.599	1.525	8,355	2.221	1.564	8,353
Help Child with Reading[2]				4.429	0.903	8,253	2.670	2.039	8,340
Help Child with Writing[2]				3.637	1.390	8,252	2.240	1.889	8,339
Help Child with Maths[2]				3.785	1.302	8,255	1.806	1.807	8,340

Notes:
[1] Coded as: 0 "Never", 1 "Less than once a month", 2 "Once a month", 3 "Once a fortnight" and 4 "Once a week."
[2] Coded as: 0 "Never", 1 "Less than once a month", 2 "1–2 times a month", 3 "1–2 times a week", 4 "3+ times a week" and 5 "every day"

A closer look at the parental investment measures reveals both expected and unexpected findings. For example, as the child ages, parents report a reduction in certain activities such as reading to the child and singing songs to them. These activities are generally associated with younger children so their decreased frequency at age 7 does not necessarily indicate a lack of engagement from parents. The decrease in other activities

[106] For clarity, these scales are collapsed to a uniform 0–5 scale for summary statistics.

is more easily explained by constraints on time rather than age appropriateness. For example, the reported frequency of engaging in physical activities and indoor games with the child falls from ages 5 to 7. This might be explained by parents having fewer hours to spend with children once the child starts formal education.[107] Most striking is the decreased incidence of help with academic activities. Intuitively, we would expect that children require increasing amounts of academic help as they progress through school; yet Table 4.10 shows that, for reading, writing and maths, the frequency of a parent helping reduces by approximately a third from ages 5 to 7.

Figure 4.10, Figure 4.11 and Figure 4.12 show the categorical distributions of the parent-reported behaviours.[108] For the majority of measures the frequency of the parenting behaviours appears to decrease over time. This could indicate that the parents are shifting to other parenting behaviours as the child ages, or that, overall, they are spending less time engaging with their children.

At age 3, as shown in Figure 4.10, none of these parent behaviours follow a normal distribution, with many of them having significant right tails. This skewed distribution indicates that most parents participate daily in the behaviours in question when the children are 3 years of age. For example, when the children are three years old, approximately half of parents report the highest frequency for reading to their children, singing songs to their children and teaching counting and numbers. The one variable that slows a slightly more differentiated pattern is teaching the child the alphabet. This indicates that this behaviour tends to vary more from one family to the other.

At age 5, as shown in Figure 4.11, there appears to be a more normal distribution of all the measured behaviours. With the exception of 'Library visits', 'Reads to child', 'Help child with reading' and 'Musical activities with child', there no longer appears to be any censoring of the variables. This indicates that these behaviours provide slightly more information about the type of interactions between the parents and their children.

At age 7, as shown in Figure 4.12, there again appears to be a substantially more normal distribution of the measured behaviours. The clear exceptions are the three variables: 'help with reading'; 'help with writing'; and 'help with maths'. For each of these variables, over 30% of the sample reports not engaging in these behaviours. Again, the histograms provide information that might have been missed by looking at summary-statistics alone.

[107]This could be linked to family income if children from high-income families spend more time in extra-curricular activities — and therefore, less time with their parents.

[108]The histograms are presented using the original response categories provided by the MCS.

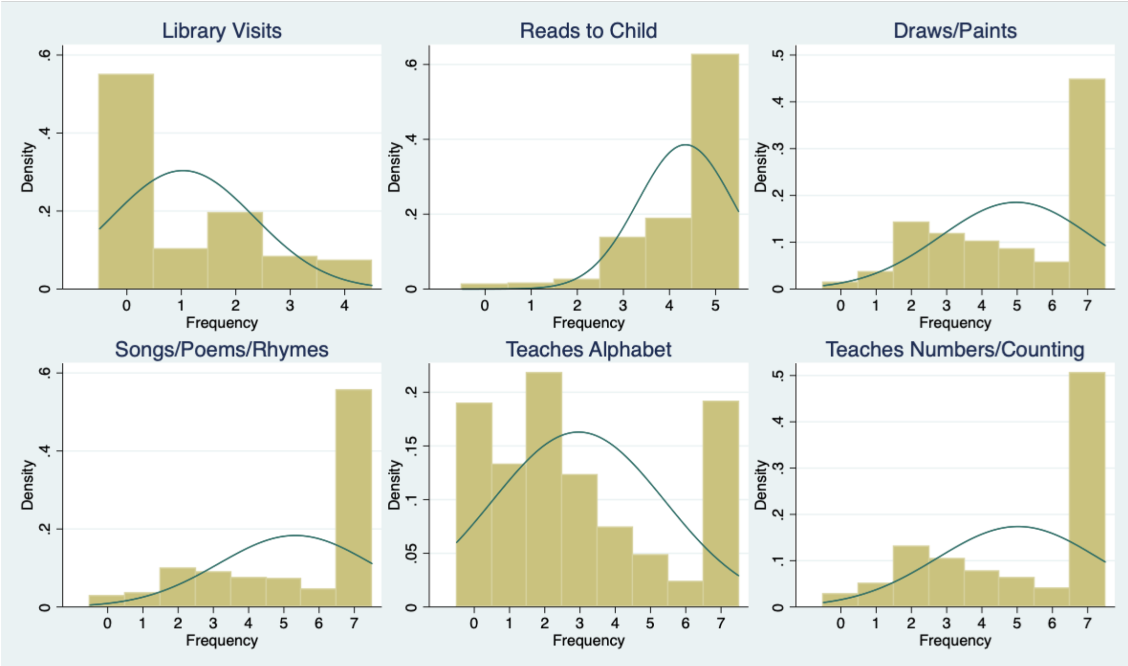


Fig. 4.10 Distribution of Parent Behaviours: MCS Sweep 2 (Age 3)

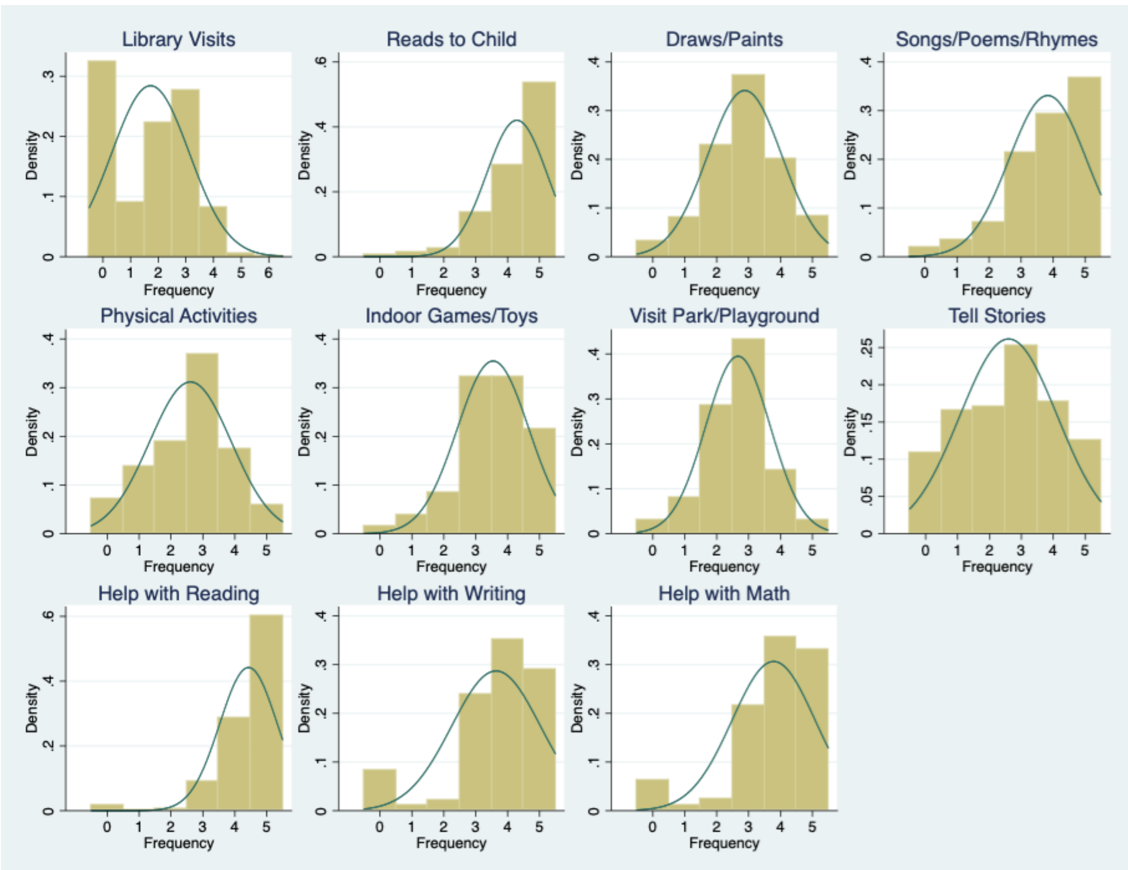


Fig. 4.11 Distribution of Parent Behaviours: MCS Sweep 3 (Age 5)

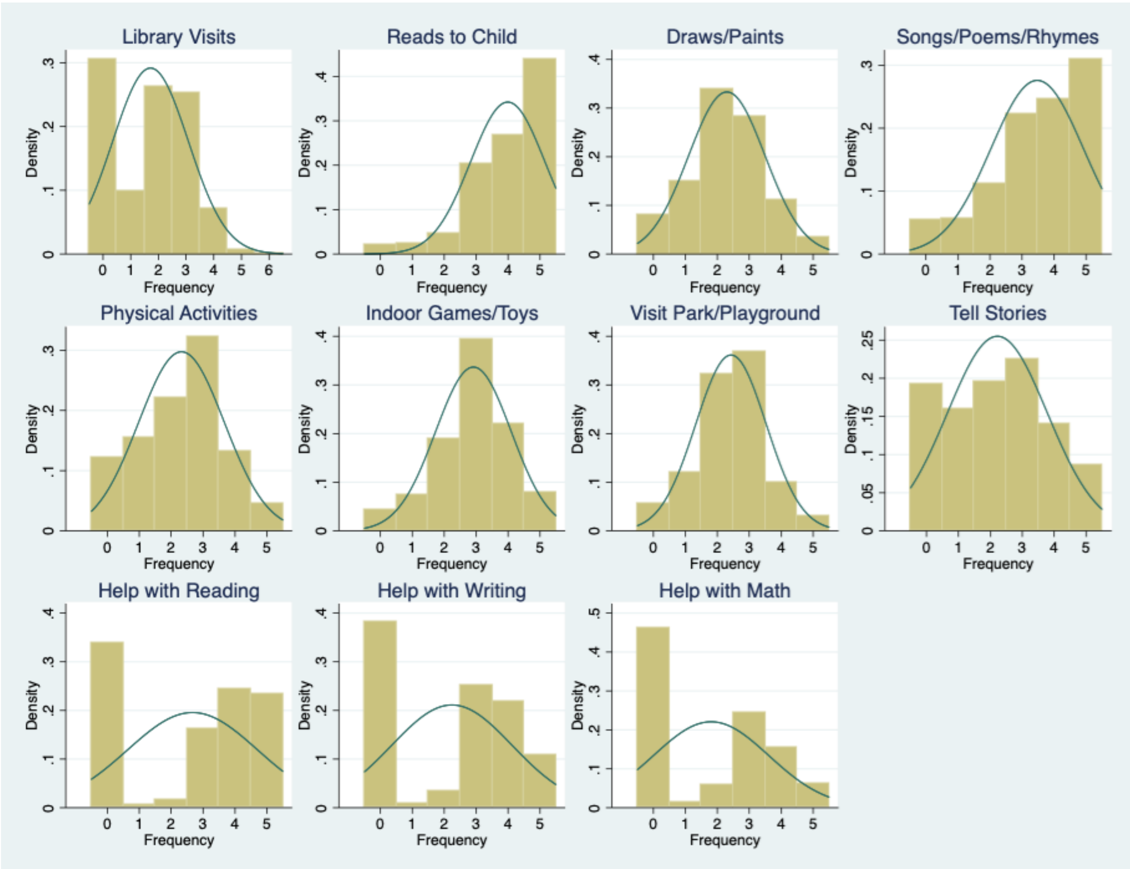


Fig. 4.12 Distribution of Parent Behaviours: MCS Sweep 4 (Age 7)

4.3 Empirical Model

The MCS data, as described above, is used to estimate the full skill formation model presented in Chapter 3. As previously discussed, this model combines a structural model and a measurement model to estimate the trajectories of skill development. In the text that follows, I briefly review the relevant equations for both the structural and measurement model before discussing them within the context of the MCS.

Structural Model

The structural model presented in Section 3.2 is a recursive model which captures the relationship between cognitive ability, non-cognitive ability and parental investment. The equations required to estimate the model are reproduced below, with an explanation of how the variables are defined in this empirical application. [109]

To review, a child's current ability θ_t is a combination of both cognitive and non-cognitive skills: $\theta_t = \left(\theta_t^C, \theta_t^N\right)'$. This set of skills in period $(t+1)$ is a function of: a child's skill in the last period — θ_t; parental behaviour (i.e. parental investment) — I_t; and observable exogenous measures of socio-economic status — X_t. Chapter 3 provides evidence that the evolution of skill over time can be expressed as the linear function

$$\theta_{t+1} = \Gamma_t \theta_t + B_t I_t + \Lambda_t X_t + \eta_t, \tag{4.1}$$

for $t \in 1, \ldots, T$, where θ_t is a latent vector of skills; I_t is an $(s \times 1)$ latent vector of parental investments[110]; X_t is an observed matrix of exogenous variables; and η_t is the error term. In the UK context, X_t contains the categorical variable measuring the household income quintile, while the latent variables: θ_t and I_t, are defined using measurement models.

All of the s types of parental investment are influenced by the matrix of observable variables X_t^I. Therefore, parental investment can be expressed as:

$$I_t = \phi_t^I X_t^I + \varsigma_t \tag{4.2}$$

where ϕ_t^I is a matrix of estimated parameters and ς_t is the error term. In this empirical application, the matrix X_t^I captures single parent household status and number of siblings.

It is assumed that a child begins life with an endowment of skill θ_0, represented by:

$$\theta_0 = \psi_0 X_0^\theta + \xi_0 \tag{4.3}$$

where ψ_0 is a matrix of estimated parameters and X_0^θ contains the period-specific measures included in X_t as well as time-invariant demographic characteristics used to capture family background, health at birth and postnatal factors which indicate early child health.

[109] For the full exposition of the structural model, the reader is directed to Section 3.2.

[110] The model explicitly allows for multiple types of parental input by defining I_t as a vector (i.e. $s \neq 1$).

In the MCS data, X_0^θ contains the measures for birth weight; mother's age at birth; and maternal education, alongside the household income quintile measure from X_t.

As described in Section 3.2 it is possible to represent the structural model using two separate linear laws of motion. Although these equations are estimated simultaneously, the two separate laws of motion are useful for describing the relevant variables in the text that follows. The linear laws of motion for non-cognitive skills and cognitive skills are

$$\theta_{t+1}^N = \gamma_{1,t}^N \theta_t^C + \gamma_{2,t}^N \theta_t^N + \beta_{1t}^N I_t^1 + \cdots + \beta_{st}^N I_t^s + \Lambda_t^N X_t + \eta_t^N, \tag{4.4}$$

and

$$\theta_{t+1}^C = \gamma_{1,t}^C \theta_t^C + \gamma_{2,t}^C \theta_t^N + \beta_{1t}^C I_t^1 + \cdots + \beta_{st}^C I_t^s + \Lambda_t^C X_t + \eta_t^C, \tag{4.5}$$

respectively, where the estimated parameters are specific to each type of skill.

Measurement Model

Fortunately, the MCS datasets contain multiple indicator variables that can be used to estimate the latent variables for the child's cognitive skills, non-cognitive skills and parental investment. The MCS-specific indicators for each latent factor are presented below, along with the relevant equations for the measurement models.

Cognitive Skills

The MCS has multiple cognitive indicators $Y_{j,t}^C, j \in \{1, \ldots, m_t^C\}$ as outlined in Table 4.5. In each period, the number of indicators is given by m_t^C with $1 \leq m_t^C \leq 3$.

This allows me to define a measurement model for cognitive ability where:

$$Y_{j,t}^C = \mu_{j,t}^C + \alpha_{j,t}^C \theta_t^C + \Phi_{j,t}^C Z_t^C + \varepsilon_{j,t}^C \tag{4.6}$$

with one equation for each of the cognitive test scores. To account for measurement error, each equation includes the matrix of covariates Z_t^C, which captures known sources of bias in the test score which are independent of the underlying latent factor. In this empirical application, Z_t^C contains the child's: age in months, ethnicity, and gender, as well as indicators for households where English is not the primary language. Each of these variables correlates with test scores, but does not directly impact cognitive ability.

By constructing Equation 4.6 for each cognitive test contained in the MCS, a system of equations can be used to estimate θ_3^C, θ_5^C and θ_7^C to include in Equation 4.5. Unfortunately, as discussed above, the MCS data on 11-year-olds only contains one applicable cognitive test, so I am unable to use a measurement model for that sweep and must include cognitive ability directly in the larger structural model. Thus, the BAS Verbal Similarities score at age 11 corresponds with θ_{11}^C in Equation 4.5 and Equation 4.4.

As the data does not allow for the use of a measurement model for this final sweep, it is assumed that there will be inherent measurement error from the cognitive measure. This measurement error can be partially corrected for by including control variables in Equation 4.5. These controls are the same measures used as covariates in the measurement models for cognitive ability in the other sweeps of the MCS. While it would be preferable to use a measurement model for all sweeps, the BAS are well validated and make it possible to extend the model to another sweep of the MCS.

Non-Cognitive Skills

Non-cognitive ability is measured in the MCS using the SDQ subscales presented in Table 4.7. Unlike cognitive ability where the number of indicator variables changes each sweep, the SDQ is the same for all sweeps and provides measures for five subscales. Therefore, $m_t^N = 5$ and the set of indicators is represented by $Y_{j,t}^N, j \in \{1, \ldots, 5\}$.

Confirming the Factor Structure: Although non-cognitive behaviour is defined in my model as a single construct, it is necessary to confirm that the chosen indicator variables are actually measuring a single construct. Fortunately, the SDQ is widely validated within the literature with the *total difficulties score* (a combined score from the subscales), often being used as a general indicator of a child's ability.

For completeness, I calculate Cronbach's alpha, α, to assess the reliability of combining the subscales in a single scale. By calculating separate values of Cronbach's alpha for multiple combined scales, with each scale removing a different subscale, it is possible to identify poorly fitting items. If the removal of any subscale significantly raises the value of α then that subscale may not identify the underlying construct. While this is unlikely due to the externally validated nature of the SDQ, it is important to confirm that the factor structure holds within the present study.

Applying the Measurement Model: After confirming that all five SDQ subscales are underlying the latent factor for non-cognitive ability, a score for the factor must be estimated. Though it would be possible to use the simple sum of the subscales, (i.e. the *total difficulties score*) it is unlikely that each measure captures the same amount of information about the true underlying level of non-cognitive skill. Instead, a measurement model is used to estimate a latent score for non-cognitive ability.

The measurement model defines a given SDQ subscale score $Y_{p,t}^N$ as:

$$(Y_{p,t}^q)^* = \mu_{p,t}^q + \alpha_{p,t}^q \theta_t^q + \Phi_{p,t}^q Z_t^q + \varepsilon_{p,t}^q \tag{4.7}$$

such that $Y_{p,t}^q = r$ if $\rho_{r-1}^p \leq (Y_{p,t}^q)^* < \rho_r^p$ where $\rho_0^p = -\infty$ and $\rho_{R_t^q}^p = \infty$. There is one equation for each of the 5 subscales. As discussed in the literature review, SDQ scores are known to differ systematically by the child's gender, ethnicity and age. There is no reason for actual non-cognitive ability to be correlated with these three factors, so they are included as Z_t^N. By constructing Equation 4.7 for all five SDQ subscales, a system of equations can be used to estimate non-cognitive ability.

Parental Investment

Unlike cognitive and non-cognitive skills, where MCS measures are used to estimate a single underlying latent factor, with parental investment there are likely to be several underlying factors that are captured by the parenting questions included in the MCS. Thus, the structural model assumes there are $S \geq 1$ types of latent parental investment.

Defining the Factor Structure: Before a measurement model can be applied, the separate types of parental investment must be identified. Table 4.9 outlines the 13 suitable measures of parental investment that I have identified in the MCS data. Applying exploratory factor analysis (EFA) to these 13 variables, I estimate the eigenvalue, RMSEA, SRMR, CFI and TLI, and use the cutoff criteria described in Section 3.3.2 to determine the number of latent investment factors to be included in the measurement model.[111]

Applying the Measurement Model: Using the $s \in \{S \geq 1\}$ latent investment factors I_t^s, identified by EFA, it is possible to create a separate measurement model for each factor. The set of MCS indicators for each factor are each expressed as $Y_{q,t}^s, q \in \{1, \ldots, m_t^s\}$ where m_t^s is the number of underlying questions for the given factor. In the MCS, parenting behaviours are reported categorically using $r \in \{1 \ldots R^q\}$ *categories*; this yields ordinal measurement models for each s factor. These take the form:

$$(Y_{q,t}^s)^* = \mu_{q,t}^s + \alpha_{q,t}^s I_t^s + \Phi_{q,t}^s Z_t^s + \varepsilon_{q,t}^s \tag{4.8}$$

such that $Y_{q,t}^s = r$ if $\rho_{r-1}^q \leq (Y_{q,t}^s)^* \leq \rho_r^q$ where $\rho_{r-1}^q = -\infty$ and $\rho_{R_t^q}^q = \infty$ with m_t^s equations, one for each parenting indicator. Each measurement equation for parenting can include a matrix of covariates Z_t^s that are known to influence the measurement of the given behaviour but are independent of the underlying parenting construct. Since no MCS variables are known to result in systematic measurement error of self-reported parental behaviour, Z_t^s is not included in the model.

Analysis Procedures

The MCS is provided as multiple data files. I use Stata 14.0 to merge these files, clean the data and calculate descriptive statistics and MPlus 8.0 (Muthén & Muthén, 2017) to conduct EFA and estimate the full dynamic model. For reference, the MPlus input file for the full dynamic model is presented in Appendix A.3. The survey weights included in the MCS files are used adjust for the MCS sampling strategy and attrition.[112] As these sample weights are based on the UK population in 2000/2001 and do not account for immigration or emigration, they may not be representative of the present-day UK population.

[111]Earlier in this dissertation, Section 3.3.1 provides a detailed explanation of EFA and the use of fit statistics. For the sake of brevity, these are not reproduced in this chapter.

[112]More information about the use of survey weights and bootstrapping are provided in Appendix A.2.

4.4 Results

Measurement Model: Cognitive Ability

Table 4.11 reports the parameter estimates for the cognitive measurement model.

Table 4.11 Measurement Model: Cognitive Ability - Parameter Estimates

PANEL A: Latent Variables

	θ_1^C	θ_2^C	θ_3^C	θ_4^C
Bracken School Readiness (age 3)	1.000 –			
BAS Naming Vocabulary (age 3)	0.962*** (0.030)			
BAS Naming Vocabulary (age 5)		1.000 –		
BAS Picture Similarity (age 5)		0.620*** (0.022)		
BAS Pattern Construction (age 5)		1.274*** (0.041)		
NFER Number Skills (age 7)			1.000 –	
BAS Word Reading (age 7)			4.758*** (0.152)	
BAS Pattern Construction (age 7)			2.379*** (0.073)	
BAS Verbal Similarities (age 11)				1.000 –

PANEL B: Observed Covariates[1]

	Age in in Months	White	Male	No English at Home	Eng./Other at Home
Bracken School Readiness (age 3)	0.135*** (0.005)	−0.008 (0.053)	−0.212*** (0.025)	−0.250*** (0.061)	−0.449*** (0.105)
BAS Naming Vocabulary (age 3)	0.091*** (0.006)	0.226*** (0.047)	−0.260*** (0.022)	−0.506*** (0.060)	−1.273*** (0.123)
BAS Naming Vocabulary (age 5)	0.067*** (0.009)	0.151*** (0.051)	−0.041 (0.027)	−0.594*** (0.068)	−1.135*** (0.101)
BAS Picture Similarity (age 5)	0.069*** (0.009)	−0.132** (0.062)	−0.083*** (0.028)	−0.079 (0.066)	−0.001 (0.111)
BAS Pattern Construction (age 5)	0.093*** (0.008)	−0.011 (0.058)	−0.164*** (0.026)	−0.213*** (0.066)	0.066 (0.112)
BAS Pattern Construction (age 7)	0.043*** (0.008)	0.091* (0.054)	−0.078*** (0.025)	−0.012 (0.066)	−0.029 (0.104)
BAS Word Reading (age 7)	0.063*** (0.008)	−0.293*** (0.053)	−0.172*** (0.026)	0.164*** (0.057)	0.185 (0.135)
NFER Number Skills (age 7)	0.063*** (0.010)	−0.104* (0.061)	0.050** (0.025)	−0.097 (0.062)	−0.261** (0.123)
BAS Verbal Similarities (age 11)	0.028*** (0.006)	−0.186*** (0.051)	0.098*** (0.025)	−0.040 (0.058)	−0.071 (0.118)

Notes:

Standard errors shown in parentheses.
*p<0.1 , **p<0.05, ***p<0.01

[1] Parameter estimates for the covariates reported as the standard deviation change in the cognitive score associated with a unit change in the covariate.

The estimates in Panel A of Table 4.11 correspond with $\alpha^C_{j,t}$ in Equation 4.6. These estimates show that at ages 3, 5, and 7 higher levels of cognitive development correspond with higher scores on all tests, as is to be expected. The factor loadings are of varying magnitudes, which indicates that each of the measures contains different amounts of information about the underlying latent variable.

As the cognitive tests vary over time, it is not possible to directly compare the magnitude of factor loadings at different ages. The relative factor loadings are still able to offer some insight into the amount of information captured by each test. At first glance, the factor loadings at age 7 are striking as the BAS scores have substantially higher loadings than the PiM score. By the nature of a measurement model, this indicates that a unit change on the two BAS tests provides more information about the underlying factor that is being measured. This likely captures the fact that the scores for the BAS take on a wider range of values than the PiM score.

The estimates in Panel B of Table 4.11 correspond with $\Phi^C_{j,t}$ in Equation 4.6. These represent observed covariates in each of the measurement equations. These coefficients show that there are significant differences in how the cognitive indicators are measured for children of the same level of skill but of differing race, gender or native language. Unsurprisingly, the child's age in months correlates with higher scores for all tests. This is in line with the information provided with the MCS that suggests using age-standardisation within any models using the MCS cognitive scores.

For other covariates the impact is less consistent. White children tend to initially score higher than their non-white peers, but over time this relationship reverses. In this sample, male respondents have lower scores than their female peers — this effect decreases over time.

In early childhood, the language spoken at home has a sizeable effect: Panel B shows that, at age 3, children who speak only English at home tend to score higher than those who speak a different language at home in addition to English and those who speak no English at home. This effect is larger for the naming vocabulary test, which is to be expected given that this test relies more heavily on language ability compared to the multi-modal tests of the Bracken School Readiness Assessment. At age 5 the impact of language ability remains sizeable for naming vocabulary but is no longer significant for picture similarity. For pattern construction, children in homes where no English is spoken score 0.213 standard deviations lower than those whose only language is English, but there is not statistically significant difference between those who speak only English and those who speak English and another language. By age 7 the impact of language spoken at home is much smaller, with children who do not speak English at home having slightly higher scores on word reading at age 7 compared to their English-speaking peers, and children from bilingual homes scoring slightly lower on the NFER.

Measurement Model: Non-Cognitive Ability

Confirming the Factor Structure

As previously discussed, the MCS uses the SDQ as a measurement of non-cognitive skills. The SDQ is a widely used measure and the Total Difficulties Score, which is a sum of the subscales has been externally validated as a single construct. To be thorough, the present analysis calculates Cronbach's alpha to confirm the reliability using a combination of the SDQ subscales as a single non-cognitive factor. Cronbach's Alpha was calculated separately for each sweep of the MCS and the results are presented below in Table 4.12.

Table 4.12 Measures of Non-Cognitive Ability: SDQ Scale Reliability

	Raw Cronbach's Alpha	Item which yields highest Cronbach's Alpha if excluded	Cronbach's Alpha if the item is excluded
MCS 2 — *Age 3*	0.655	Emotional Symptoms	0.643
MCS 3 — *Age 5*	0.681	Emotional Symptoms	0.662
MCS 4 — *Age 7*	0.713	Pro-Social Behaviour	0.694
MCS 5 — *Age 11*	0.736	Pro-Social Behaviour	0.727

Notes:
- Cronbach's alpha was calculated for the scale omitting each of the included variables.
- Only the summarised results are presented above. Full results can be provided upon request.

In Table 4.12 it can be seen that for all four sweeps, the removal of any of the subscales results in a lower value for α. This indicates that there is no SDQ subscale that reduces the reliability of the full SDQ scale. To confirm this finding, EFA was conducted for the SDQ measures. Though the full findings are not reported in this dissertation, they point to a single factor solution with statistically significant loadings on all of the SDQ subscales. The results from these two assessments of fit led to the decision to keep all of the SDQ scales for the non-cognitive factor score.

Results from the Measurement Model:

Table 4.13 reports parameter estimates for the non-cognitive measurement model. Panel A shows the factor loadings on each indicator which correspond with $\alpha^N_{j,t}$ in Equation 4.7. Because the SDQ measures correspond with the presence of negative behaviours, the factor loadings are set as negative values. This makes the latent factor for non-cognitive ability easier to interpret. At ages 3, 5, and 7, higher levels of non-cognitive ability correspond with lower parental reports of hyperactivity, emotional problems, peer problems, and conduct problems, but higher levels of pro-social behaviour. The factor loadings vary in magnitude, indicating that each of the measures contains different amounts of information about the underlying latent variable.

Table 4.13 Measurement Model: Non-Cognitive Ability - Parameter Estimates

Panel A:	Latent Variables			
	θ_1^N	θ_2^N	θ_3^N	θ_4^N
Peer Problems	–1.000	–1.000	–1.000	–1.000
	–	–	–	–
Pro-Social Behaviour	–1.171***	–1.232***	–1.112***	–0.863***
	(0.054)	(0.056)	(0.044)	(0.033)
Conduct Problems	–1.975***	–1.575***	–1.412***	–1.215***
	(0.081)	(0.056)	(0.049)	(0.036)
Emotional Symptoms	–0.808***	–1.152***	–1.171***	–1.266***
	(0.035)	(0.046)	(0.045)	(0.040)
Hyperactivity	–2.049***	–2.520***	–2.373***	–2.030***
	(0.083)	(0.091)	(0.085)	(0.062)

Panel B:	Observed Covariates[1]							
	MCS 2		MCS 3		MCS 4		MCS 5	
	Age 3		*Age 5*		*Age 7*		*Age 11*	
	White	Male	White	Male	White	Male	White	Male
Peer Problems	–0.178***	0.152***	–0.162***	0.116***	–0.172***	0.125***	–0.051	0.125***
	(0.050)	(0.027)	(0.056)	(0.028)	(0.052)	(0.025)	(0.063)	(0.027)
Pro-Social Behaviour	0.092*	0.231***	0.079	0.282***	0.058	0.329***	–0.025	0.346***
	(0.049)	(0.025)	(0.057)	(0.026)	(0.060)	(0.024)	(0.053)	(0.027)
Conduct Problems	0.055	0.099***	0.025	0.212***	0.161***	0.228***	0.219***	0.209***
	(0.056)	(0.025)	(0.057)	(0.024)	(0.062)	(0.025)	(0.058)	(0.026)
Emotional Symptoms	–0.052	0.000	–0.044	–0.043	–0.047	–0.024	0.077	–0.080***
	(0.064)	(0.027)	(0.057)	(0.027)	(0.057)	(0.026)	(0.053)	(0.028)
Hyperactivity	0.011	0.233***	–0.002	0.292***	0.061	0.353***	0.094*	0.378***
	(0.067)	(0.028)	(0.057)	(0.026)	(0.059)	(0.024)	(0.049)	(0.024)

Notes:
- Standard errors shown in parentheses.
- *p<0.1 , **p<0.05, ***p<0.01

[1] Parameter estimates reported as the standard deviation change in the latent cognitive ability for a unit change in the covariate.

There are several notable findings in this table. First, the hyperactivity measure has the highest factor loadings in all four periods with conduct problems being the second highest in the first three periods. This pattern indicates that a higher score in conduct problems or hyperactivity scale will have a larger negative impact on the latent measure as compared to a higher score in the other categories. Finally, there is variation in the loadings across periods, which shows that although the questions used to generate the measure were consistent over time, certain behaviours have a larger influence on the generated factors during different stages of childhood. This might show that, for example, there is less variation in observed pro-social behaviour by age 11, so it is a less indicative measure of underlying ability, compared to age 3 where there is higher variation.

The estimates in Panel B of Table 4.13 correspond with $\Phi_{j,t}^N$ in Equation 4.7. These represent observed covariates in each of the measurement equations. These coefficients show that there are significant differences in how the non-cognitive indicators are measured for children of the same level of ability but of differing race or gender. White children tend to have higher reported levels of conduct problems at ages 7 and 11 and lower reported problems with their peers at ages 3, 5 and 7. Though significant, the effect size of this measurement error for race is small, accounting for no more than 0.068 of a standard deviation compared to non-white children. Across all ages, boys have higher reported levels of hyperactive behaviour, conduct problems and peer problems and lower reported levels of pro-social behaviour than girls with the same underlying non-cognitive skill. The magnitude of this difference in the perceived behaviour of boys compared with perceived behaviour of girls grows over time; at age 11, being male corresponds with having nearly a fifth of a standard deviation higher score on the hyperactivity and conduct measures.

Measurement Model: Parental Investment

Determining the Factor Structure

In Section 4.2, I discussed the multitude of measures included in the MCS that capture parental input, with Table 4.9 outlining those included in the present analysis. EFA allows me to determine the number of parenting constructs captured by these indicators and to identify which variables load onto each factor. EFA was conducted for each sweep of the MCS and the measures of fit are presented below in Table 4.14.

Table 4.14 Parental Investment Measures: EFA Measures of Fit

	Eigenvalue	RMSEA	CFI	TLI	SRMR
Age 3 – MCS Sweep 2					
One Factor	2.218	0.083	0.884	0.807	0.078
Two Factor	1.190	0.044	0.986	0.946	0.024
Three Factor	0.862	0.000	1.000	1.000	0.000
Age 5 – MCS Sweep 3					
One Factor	3.551	0.086	0.835	0.793	0.081
Two Factor	1.334	0.060	0.937	0.898	0.048
Three Factor	1.103	0.032	0.987	0.971	0.020
Four Factor	0.857	0.028	0.993	0.978	0.014
Age 7 – MCS Sweep 5					
One Factor	3.324	0.117	0.754	0.692	0.121
Two Factor	1.720	0.041	0.977	0.963	0.032
Three Factor	1.017	0.040	0.983	0.963	0.024
Four Factor	0.907	0.029	0.994	0.981	0.014

Notes:
RMSEA — root mean square error of approximation. CFI — comparative fit index.
TLI — Tucker-Lewis index. SRMR — standardised root mean square residual.

The relevant diagnostic measures point to two factors of parental investment at age 3, and three factors at ages 5 and 7. As discussed above, assessment of fit using EFA is based on a set of fit-statistics. At age 3, for a two-factor model, the reported EFA is under 0.05 with CFI above 0.95, TLI above 0.90 and SRMR well below 0.08.[113] At age 5, for the three factor model, the reported RMSEA is significantly under 0.05; CFI and TLI are above 0.90; and SRMR is below 0.08. The results for age 7 are slightly less straightforward: though there are three factors with eigenvalues greater than one, the two-factor model satisfies the relevant fit-criteria. As the eigenvalues and fit-statistics support either two or three factors, deciding how many factors to retain requires an inspection of the factor structure. A comparison of the factor structure of the two- and three-factor models finds that the three-factor model is consistent across ages 5 and 7.

The results from the EFA confirm the findings of Hernández-Alava and Popli (2017), that there is more than one latent variable underlying parental investment. As the choice

[113]For the cutoff points used for these fit-statistics, the reader is directed to the detailed explanation of EFA presented in Section 3.3.2.

of parental investment indicators differs from Hernández-Alava and Popli (2017), the resulting investment factors do not match those presented in their work.[114]

Table 4.15 contains the factor structure identified using EFA for the three factor model. For each variable, 'X' marks the factor with the largest significant factor loading.

Table 4.15 Parental Investment Measures: Factor Structure

	MCS 2 *Age 3*		MCS 3 *Age 5*			MCS 4 *Age 7*		
	I_1^1	I_1^2	I_2^1	I_2^2	I_2^3	I_3^1	I_3^2	I_3^3
Child visits the Library	X		X			X		
Reads to Child	X		X			X		
Draws/Paints with Child		X		X			X	
Songs/Poems/Rhymes with Child		X		X			X	
Helps Child Learn Alphabet		X						
Teaches Child Counting		X						
Physical Activities with Child				X			X	
Indoor Games/Toys with Child				X			X	
Take Child to Park/Playground				X			X	
Tell Stories to Child				X			X	
Help Child with Reading			X					X
Help Child with Writing					X			X
Help Child with Maths					X			X

Notes:
[1] For ease of exposition, the estimated factor loadings are not included in this dissertation.
[2] Analysis was conducted using MCS provided survey weights.

When the children are 3 years old, the EFA points towards two latent factors in parental investment. The first factor is based on the frequency of 'reads to child' and 'child visits the library', while the second factor is based on the remaining measures. When the child is 5 and 7 years old, the analysis points to three factors: the first again corresponding to reading and library visits; the second being informal activities between the mother and child; and the third representing activities relating specifically to schoolwork. I have labelled these three parenting factors: *literacy activities* (I_t^1); *parent child interactions*, (I_t^2); and *academic activities* (I_t^3).

For the majority of indicators, the factor structure is identical between ages 5 and 7. The only exception is the measure "helps the child with reading", for which the EFA results indicated significant loadings on both the first and third factor, with the highest loading changing from ages 5 to 7. For this dissertation, I have opted to maintain consistency between periods and this factor was included in the measurement equation for the third parental investment factor.

[114] The present analysis focuses on behaviours requiring parent-child interaction and includes five additional measures of parent-child interaction that were not included by Hernández-Alava and Popli (2017) while omitting measures of 'consistent bedtime' and 'time spent watching television'.

Results from the Measurement Model

Table 4.16 presents the parameter estimates that correspond with $\alpha^s_{q,t}$ in the measurement model represented by Equation 4.8. Though the model allows for covariates, no observed covariates were consistently correlated across the measurement model for parental investment, so Z^s_t is omitted from the equation for the analysis.

The first latent variable for each period is constructed using measures of 'reads to child' and 'visits the library'. The coefficients on these two factors shift over time with the factor loading on 'visits the library' increasing over time from 40.6% of the loading on 'reads to child' at age 3, compared to 120.6% at age 7. This change in the loading over time could indicate that for older children, the measurement for trips to the library provides more information about the underlying behaviour of parents. This assumption is in line with the more balanced distribution of the 'visits the library' measure in children ages 5 and 7, compared to the left skew seen at age 3 where a large portion of the sample did not report any library visits. Thus, at age 3, the lack of variability means that the measure is less able to tell us about the variation in parenting within the sample.

For age 3, the second factor contains the remaining measures of parent-child interaction. The highest factor loading corresponds with 'teaches child counting'. At ages 5 and 7, the second latent variable contains 'non-academic' parent-child interactions. In both periods, the highest loading corresponds with 'playing indoors with the child'. Though the factor loadings change slightly from ages 5 to 7 the order of magnitudes remains the same, indicating a consistent relationship between the measures over time.

The final latent variable examined parenting behaviours directly relating to academic work. The factor loadings between ages 5 and 7 were very consistent for this latent variable. For both ages 5 and 7, 'help with writing' had the highest loading.

Table 4.16 Measurement Model: Parental Investment - Parameter Estimates

	MCS 2 *Age 3*		MCS 3 *Age 5*			MCS 4 *Age 7*		
	I_1^1	I_1^2	I_2^1	I_2^2	I_2^3	I_3^1	I_3^2	I_3^3
Literacy Activities								
Reads to Child	1.000		1.000			1.000		
	–		–			–		
Visits the Library	0.406		0.656			1.206		
	(0.075)		(0.138)			(0.287)		
Parent Child Interactions								
Songs/Rhymes with Child		1.000		1.000			1.000	
		–		–			–	
Draws/Paints with Child		0.583		1.259			1.415	
		(0.027)		(0.043)			(0.047)	
Helps Child Learn Alphabet		0.841						
		(0.029)						
Teaches Child Counting		1.166						
		(0.050)						
Physical Activities/Games				1.220			1.390	
				(0.042)			(0.047)	
Indoor Games/Toys				1.306			1.563	
				(0.040)			(0.051)	
Take Child to Park				0.800			0.883	
				(0.037)			(0.034)	
Tell Stories to Child				1.004			1.095	
				(0.033)			(0.037)	
Academic Activities								
Help with Maths					1.000			1.000
					–			–
Help with Writing					1.124			1.159
					(0.047)			(0.017)
Help with Reading					0.822			0.957
					(0.025)			(0.014)

Notes:

p<0.01 for all values
Standard errors shown in parentheses.

Structural Model:

The parameter estimates of the structural model are presented in Table 4.17. The parameters reported in Panel A correspond with the coefficients Γ_t and B_t in Equation 4.1. Panel B reports Λ_t, representing the impact of the covariates X_t on latent ability.

Table 4.17 Structural Model: Parameter Estimates

PANEL A: Latent Factors[1]

	Cognitive Ability (θ^C_{t+1})			Non-Cognitive Ability (θ^N_{t+1})		
	θ^C_5	θ^C_7	θ^C_{11}	θ^N_5	θ^N_7	θ^N_{11}
Lagged Ability:						
Cognitive (θ^C_t)	0.731***	0.901***	0.458***	0.176***	−0.016	0.061***
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)
Non-Cognitive (θ^N_t)	0.122***	0.054***	0.063***	0.773***	0.874***	0.820***
	(0.016)	(0.017)	(0.014)	(0.011)	(0.012)	(0.013)
Lagged Parental Inputs:						
Literacy Activities (I^1_t)	0.200***	0.136***	0.086***	0.185***	0.197***	0.175***
	(0.023)	(0.025)	(0.027)	(0.022)	(0.025)	(0.035)
Parent-Child Interaction (I^2_t)	0.109***	0.030*	0.017	0.113***	0.152***	0.094***
	(0.017)	(0.016)	(0.014)	(0.016)	(0.016)	(0.016)
Academic Activities (I^3_t)		0.011	−0.056***		0.147***	−0.001
		(0.017)	(0.014)		(0.018)	(0.015)

PANEL B: Observed Covariates[2]

	Cognitive Ability (θ^C_{t+1})			Non-Cognitive Ability (θ^N_{t+1})		
	θ^C_5	θ^C_7	θ^C_{11}	θ^N_5	θ^N_7	θ^N_{11}
Second Income Quintile	0.030	0.101*	0.092**	0.023	0.034	0.049
	(0.055)	(0.061)	(0.045)	(0.062)	(0.052)	(0.058)
Third Income Quintile	0.112*	0.126*	0.134***	0.106	0.182***	0.117*
	(0.062)	(0.066)	(0.046)	(0.065)	(0.060)	(0.063)
Fourth Income Quintile	0.105	0.095	0.102*	0.107	0.194***	0.161**
	(0.065)	(0.071)	(0.057)	(0.077)	(0.074)	(0.067)
Fifth Income Quintile	0.090	0.177**	0.087	0.155*	0.271***	0.214***
	(0.d067)	(0.086)	(0.064)	(0.088)	(0.080)	(0.075)
Month of Birth	0.007	−0.002	0.001	0.005	0.004	−0.003
	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)

Notes:
- All models are estimated using provided survey weights.
- *p<0.1 , **p<0.05, ***p<0.01
- Standard errors in parentheses.

[1] Parameter estimates for the latent factors reported as the standard deviation change in the latent score associated with a standard deviation change in the lagged latent score.

[2] Parameter estimates for the observed covariates reported as the standard deviation change in the latent score associated with a unit change in the covariate.

For both cognitive and non-cognitive ability, higher levels of the ability in one period predict higher levels of ability in the next period. This is evidence of the theory that skills are self-productive. This autoregressive effect is statistically significant over all three periods of measurement. For cognitive ability, the estimate of self-productivity decreases in magnitude for the final period. This could be caused by attenuation bias. Alternatively, it could be because θ^C_{11} was directly measured, as opposed to cognitive ability in other periods which was constructed using a measurement model. The self-productivity of non-cognitive ability is similar in magnitude across time. As the indicators used for the measurement model were consistent, the consistency of the parameter estimate tells us that the underlying self-productivity of skills is consistent over time.

For cognitive ability there is evidence of cross-productivity of non-cognitive ability with the coefficients being statistically significant at the 1% level in all periods. This indicates that children with higher non-cognitive skills go on to have higher cognitive skills in the next period. The magnitude of this estimate is halved from the first to second period, indicating that cross-productivity is largest in early childhood. For non-cognitive development, cross productivity is also seen at ages 5 and 11. Strikingly, cross productivity is not observed at age 7. Again, the estimate of cross-productivity is lower in the last period, indicating that the effect of cross-productivity fades over time.

The three latent variables for parental investment are all significant determinants of cognitive and non-cognitive skills at some stage of development. By examining the periods where investment is a significant predictor of latent ability, it is possible to identify sensitive periods of investment. Trends in the magnitude of this coefficient over time might indicate when investment is most effective in improving underlying ability.

The first investment factor, which was constructed using measures of how often the parent read to the child and took them to the library, is statistically significant across all periods for both cognitive and non-cognitive ability. Not only is this factor significant in all periods, but it also has the largest effect size. For cognitive ability, the magnitude of this parameter falls over time; one standard deviation increase in literacy activities corresponds with a 0.200 standard deviation increase in cognitive ability at age 5, but only a 0.086 standard deviation increase at age 11. This points to children's cognitive ability being particularly sensitive to this type of investment in pre-primary years, but that at later stages the effect on cognitive ability diminishes.

While it is unsurprising that literacy activities correspond with higher cognitive ability, this factor is also significant for non-cognitive ability at all ages, which illustrates that the time spent on these activities extends benefits beyond literacy. This finding is particularly notable as some parents might think that older children no longer benefit from this type of activity, when in fact it has benefits that are related to non-cognitive skills. The magnitude of the estimated parameter for this latent investment factor is relatively consistent over time, with the largest estimate of 0.197 corresponding with θ_7^N. This suggests that parents who are one standard deviation above the mean correspond with children having 0.197 standard deviation higher level of non-cognitive ability.

The second investment factor is significant in all periods for non-cognitive ability but the magnitude decreases over time. This factor is constructed using non-academic parent-child interactions. The magnitude of these coefficients is approximately half to three-quarters of those seen on the first factor, demonstrating that this type of parental investment has smaller effects on non-cognitive ability. This factor is only significantly related to cognitive ability at age 5, though the coefficient is approximately 55% the magnitude of that on literacy activities. This aligns with past research which has shown that parental investment has a greater impact on early cognitive ability and this effect diminishes over time.

The third investment factor, constructed using activities involving academic interactions, delivers the most surprising results. More specifically, this type of parental investment has a small but statistically significant negative coefficient on age 11 cognitive ability. It seems counter-intuitive that more time spent on homework at age 7 corresponds with lower ability at age 11, but this might indicate reverse causality, with parents at age 7 responding to children who are struggling academically by providing more guidance.

In addition to the small impact on cognitive ability, the factor of academic activities has a fairly large predictive power for non-cognitive ability at age 7, but is insignificant in the other periods. The magnitude of this coefficient on non-cognitive ability is similar to that on the parent-child interaction factor. These results might be capturing the behavioural benefits of this time spent with the child and indicate that children ages 5 to 7 are particularly sensitive to one-on-one time spent on quiet activities. Perhaps this time involves more practice of behavioural skills and reinforcement of these behaviours.

Panel B of Table 4.17 presents the parameter estimates for Λ_t in the structural model shown in Equation 4.1. This parameter measures the effect of X_t (family income) on latent ability, separate from other measures of parental investment. Compared with the lowest income quintile, all of the other income quintiles have either positive or non-significant co-efficients. For the coefficients that are significant, the magnitude of the estimates for Λ_t are similar in magnitude to those for all three types of parental investment with values ranging from 0.102 to 0.271. This finding taken on its own is not surprising as there is substantial evidence in the literature that children from higher income households score higher on various measures of cognitive and non-cognitive ability. It is notable that the parameter estimates for Λ_t for cognitive ability are not all significant, as usually children in higher income quintiles perform better on measures of cognitive ability. Within the context of the model, any significant parameters for Λ_t provide further support of the model structure which separates family income from parental investment. As family income is a significant predictor of latent ability on its own, any model including family income as an underlying measure of investment would likely vastly overstate the effect of parental behaviour.

Table 4.18 presents the parameter estimates for the covariates for initial ability. These correspond with ξ_0 in Equation 4.3. Higher levels of both cognitive and non-cognitive ability are associated with increased family income and maternal education.

Table 4.18 Structural Model: Parameter Estimates Initial Period Covariates

	Cognitive Ability (θ_0^C)	Non-Cognitive Ability (θ_0^N)
Second Income Quintile	0.110*	0.101*
	(0.058)	(0.055)
Third Income Quintile	0.304***	0.200***
	(0.058)	(0.064)
Fourth Income Quintile	0.380***	0.255***
	(0.062)	(0.062)
Fifth Income Quintile	0.520***	0.205***
	(0.062)	(0.069)
Birthweight (kg)	0.200***	0.098***
	(0.027)	(0.026)
Mother's Age at Birth	0.008**	0.019***
	(0.003)	(0.003)
Higher Degree	0.977***	0.366***
	(0.093)	(0.093)
First Degree	0.776***	0.519***
	(0.063)	(0.077)
Post-Grad. Dipl. & Cert.	0.465***	0.383***
	(0.064)	(0.068)
A/AS/A Levels	0.424***	0.473***
	(0.071)	(0.070)
O-Level/GCSE (grades A–C)	0.257***	0.350***
	(0.056)	(0.056)
O-Level/GCSE (grade <C)	−0.003	0.078
	(0.074)	(0.064)
Other Qual. (Inc. Overseas)	0.158	0.409***
	(0.132)	(0.110)
Month of Birth	−0.001	−0.004
	(0.005)	(0.004)

Notes:
- Structural model estimated using provided survey weights.
- *p<0.1 , **p<0.05, ***p<0.01
- All estimates provided in partially standardised form. The reported coefficients indicate the change in standard deviation of the outcome corresponding to a unit change in the parameter listed.
- Standard errors in parentheses.

Unsurprisingly, higher family income corresponds with significantly higher estimates of latent cognitive ability at age 3, with children in the top quintile of the income distribution having latent factor scores 0.520 standard deviations above the mean. Non-cognitive ability at age 3 is also correlated with family income, though the effect is smaller, with the estimated latent scores for children in the top quintile being 0.205 standard deviations above the mean.

Mother's age at the time of her child's birth has limited predictive power as each additional year only corresponds with an 0.008 increase in cognitive ability and an 0.019 increase in non-cognitive ability. Birthweight appears to have a large predictive effect with coefficients of 0.200 and 0.098 for cognitive and non-cognitive ability respectively. This coefficient is slightly misleading as birthweight is included in the model in kilograms. As the standard deviation is only 0.575 kg, a single unit decrease in birthweight corresponds with a child being nearly 2 standard deviations below the mean.

Finally, maternal education is highly predictive of both types of ability, with the effect size increasing as the level of educational attainment rises. At age 3, the cognitive ability of children whose mothers hold higher degrees is 0.977 standard deviations above the mean compared to children whose mothers hold no qualifications. The effect on non-cognitive ability appears to vary less by qualification: the coefficients on the indicators of A-levels, post-graduate diplomas, first degrees and higher degrees are of similar magnitudes. Having a mother with any of these qualifications corresponds with non-cognitive ability approximately 0.4 to 0.5 standard deviations higher than for those whose mothers have no qualifications.

Table 4.19 provides the estimated coefficients for ϕ_t^I in Equation 4.2. For most types of investment, the coefficient on single parent status is not significant. The only exception to this is parent-child interactions at age 7, which correspond with children in single parent households experiencing a frequency of these interactions 0.148 standard deviations above those in dual-parent households.

Table 4.19 Structural Model: Parameter Estimates Covariates for Investment

	Age 3		Age 5		Age 7	
	No. of Siblings	Single Parent	No. of Siblings	Single Parent	No. of Siblings	Single Parent
Literacy Activities (I_t^1)	−0.143***	−0.045	0.002	−0.205	−0.322***	0.053
	(0.051)	(0.101)	(0.090)	(0.129)	(0.089)	(0.106)
Parent–Child Interaction (I_t^2)	−0.037	0.011	−0.096**	−0.059	−0.027	0.148**
	(0.035)	(0.074)	(0.047)	(0.069)	(0.046)	(0.062)
Academic Activities (I_t^3)			−0.006	−0.006	−0.073*	−0.030
			(0.044)	(0.074)	(0.039)	(0.065)

Notes:
- Structural model estimated using provided survey weights.
- *p<0.1 , **p<0.05, ***p<0.01
- Standard errors in parentheses.
- Estimates provided in partially standardised form. Coefficients indicate the change in standard deviation of the outcome corresponding with a unit change in the listed covariate (e.g. No. of siblings).

The impact of family size on the latent parenting factors varies between the specific factors and time periods. For the periods in which the number of siblings is significant, the coefficients are all negative. This means that as family size increases, the amount of time spent in these activities with each child decreases. The largest impact of family size is seen at age 7, with each additional child corresponding to 32.2% of a standard deviation lower levels of literacy activities.

At first glance, the fluctuating significance of these co-variates may raise doubts as to the validity of the model. However, when considered within the entire structure of the model, there is a simple explanation for this change in significance: the underlying factor structure of the latent investment measures can vary over time and the factor loadings tend to change in relative magnitude between periods. For example, the factor loadings on ‘literacy activities’ fluctuate from period to period: ‘reads to child’ has a factor loading of 1 in each period, while the ‘library visits’ factor has a relatively low factor loading of 0.656 at age 5 — low compared to its loading of 1.206 at age 7. As the impact of family size differs for each of the underlying indicators, the changing factor structure is able to explain why the significance of this variable changes over time. Similar explanations hold for the other co-variates. This is confirmed by examining the correlation table which shows relatively consistent relationships between the individual factors and co-variates over time.

The results above show that there are specific sensitive periods for parental investment, and that cognitive and non-cognitive ability show signs of cross- and self-productivity. As in past research, the findings of this analysis show that non-cognitive ability is generally more sensitive to parental investment in later stages whereas cognitive ability has the highest effect size at the pre-primary stage.

4.5 Discussion

By applying my updated version of the skill formation model to longitudinal data from the MCS, this chapter satisfies two key aims of this dissertation. First, it confirms the suitability of this modified model to estimate the trajectories of skill development: this proves that the model is fit-for-purpose. Second, the empirical application provides detailed estimates of the role that different types of parental input play in skill formation in the UK: these estimates were one of the primary goals of this dissertation. Below, I discuss how this chapter meets these two objectives, and then explore how this chapter fits within this dissertation as well as within the existing literature.

Support for Methodological Approach

As explained in Chapter 3, this dissertation presents an adaptation of an existing methodology for modelling skill development in children. While the empirical findings and theoretical literature justification for this approach are presented in earlier parts of this dissertation, it is necessary to prove that the methodology I propose works in practice. Therefore, this chapter serves to demonstrate how the empirical model introduced in Chapter 3 can be applied to existing data and to show how this new approach yields estimates which provide information not captured by the existing methodologies.

Since it builds on an existing theoretical understanding, this empirical application also adds to the growing body of evidence in support of the theory of skill formation presented by Cunha and Heckman (2008) and the estimation of the model provided by Cunha et al. (2010). Specifically, the original methodology does not differentiate between the financial resources of the family and investment in the form of parenting behaviour. While their estimates provide valuable insight into the patterns of skill development seen in US children, my analysis yields new insight into how specific parenting factors drive the sensitive periods of investments which Cunha and Heckman identify.

I am not the first to identify the need for different types of parental investment and recently Hernández-Alava and Popli (2017) have estimated skill formation using a model which separates parenting from family resources. Unfortunately, their analysis was limited by the data available at the time of their study, as well as by the lack of detail presented in their justification for modelling parental input in this way. The analysis presented in this chapter bridges the gap between these two models: it allows for the identification of multiple types of parental investment; it examines the specific impacts from these different types of investment; and it captures the role that family resources play over and above parenting behaviour. The results prove the suitability of the model for measuring the relationship between parenting behaviour and childhood ability. Specific findings will be discussed more in the following section.

Empirical Evidence

In addition to providing proof of concept for my empirical methodology, this chapter also provides UK specific estimates of skill formation. These estimates are especially valuable as I have applied a skill formation model to a nationally representative sample which includes female and minority respondents — this in contrast to the original work of Cunha and Heckman (2008) which was limited to a small sample of American males.

I find that cognitive ability and non-cognitive ability are both strongly persistent over time. Furthermore, early non-cognitive skills are strong determinants of cognitive ability, indicating that there is a significant cross-productive effect of non-cognitive ability in the early formation of cognitive ability. Early cognitive ability is also predictive of later non-cognitive ability, indicating that non-cognitive skills are reliant on the presence of cognitive ability for their development. This cross-productivity is particularly interesting as it suggests the potential for intervention targeting children with poor initial cognitive or non-cognitive performance to yield improvements across the board in future development.

The estimates from this model not only measure the relationship between parenting behaviour and skill development, but also serve to strengthen our understanding of how parenting indicators can be used to capture underlying constructs. Put differently, by examining the available parenting measures in the MCS, I was able to identify a set of parenting indicators that capture various ways that parents spend time with their children. Using EFA, I identified three unique parenting factors: literacy activities, parent-child interactions and academic activities. By including these three activities as inputs in the production function for childhood ability, I was able to provide updated estimates on skill formation in primary school children in the UK.

In line with existing research, the results from the present study show that parental input has a significant influence on both cognitive and non-cognitive skills in the early stages of development but that for older children, parental input is a less important determinant of both types of skills. In general, the impact of parental investment is larger for non-cognitive ability, indicating that parenting behaviours might have more capacity to change the outcomes in this area of development.

For the purpose of this study, there is a specific focus on parental inputs that are not reliant on access to resources. The fact that these low- or no-cost behaviours are significant determinants of skill highlights the ability for families across the SES spectrum to exert positive developmental changes on their children. This is especially true at the early stages of development where increases in parental investment have approximately a third of the predictive power of existing cognitive ability.

On a more detailed level, it can be seen that there are different types of parental investment, and that it is not sufficient to treat all of these types of parental investment as the same. Past research has assumed that all parenting behaviours are measuring the same underlying factor, but this dissertation highlights the presence of multiple underlying types of investment. Results from exploratory factor analysis indicate that the measures of investment in the MCS fall into three categories. This dissertation defines these as ‘early literacy activities’, ‘parent-child interaction’ and ‘academic activities’. If these are treated as three separate factors, it is possible to see that there are differing periods of sensitivity to different parental behaviours.

This distinction in categories is particularly relevant for providing policy recommendations to parents. While past work has highlighted that early investment is most productive in cognitive development, but that non-cognitive development continues to benefit from investment in later ages (Cunha & Heckman, 2008; Cunha et al., 2010), the findings from this analysis show that there is a more nuanced explanation with specific types of parenting behaviour showing benefits at all stages. Most notable is that activities that might appear to be ‘non-academic’ in nature show reasonable effect sizes for cognitive development at certain stages. Similarly, at age 11 there is evidence that ‘academic interactions’ between children and parents are correlated with sizeable decreases in the presence of problematic behaviour. This finding provides evidence to encourage parents to engage with their child’s academic work even in cases where it might not seem necessary.

Though these findings provide valuable insight into the nature of skill development they should be taken with several notes of caution. First, despite using longitudinal data, the model is still correlational in nature and, as discussed earlier, there is potential for reverse causality or spurious correlation. Secondly, though the MCS is a widely used and validated dataset, the present study is limited by the types of questions included in the survey. Thirdly, the model is not designed to capture genetic effects which may contribute to child development. The data used in this analysis does not allow me to explore the role of genetics in skill development directly, but some aspects are captured by proxies in the data such as parental education. Finally, the model is dependent on the structural assumptions as discussed in the methodology section: if any of those were to be violated the results would be less robust.

Even with these potential limitations, the findings have direct policy implications. By identifying how sensitive periods of development vary by type of parental investment, I am able to provide more precise recommendations for what types of investment are most effective at each age. While these findings are important, further work needs to be done that includes measures of parental attitudes to understand the mechanisms driving effective parental behaviours. In the meantime, my findings do suggest that, before age 5, the time parents spend with their children can have significant effects on cognitive and non-cognitive ability. This time need not be formally structured or focused on specific academic outcomes, but can be as simple as arts and crafts, or reciting nursery rhymes.

CHAPTER 5

Empirical Application II: Canada

The final substantive chapter of my PhD applies the methodology discussed in Chapter 3 to Canadian data. This provides further proof of the strength of the model as well as an international context for the UK analysis discussed in Chapter 4. This chapter is organised in the following manner. To begin, Section 5.1 contextualises the present empirical application, with a discussion of the existing Canadian research on parenting and a brief overview of how Canadian policy differs from that in the UK. Next, Section 5.2 introduces the data used in this empirical application, examines the variables chosen for this analysis and provides at descriptive statistics for the survey sample and relevant measures. Building on this introduction to the data, Section 5.3 discusses the methodological considerations taken to adjust the model to the Canadian data. The results obtained using this methodology are presented and explained in Section 5.4. The final section, Section 5.5, provides concluding remarks on this empirical work and its importance for policy.

5.1 Introduction

Using a longitudinal Canadian dataset, this chapter presents further evidence to support the use of an empirical methodology which differentiates between parental behaviours and family resources when modelling parental investment. Though Chapter 4 has shown the suitability of my proposed methodological framework for estimating skill formation using data from the UK, it is important to see how this modelling technique applies to other contexts. The present application not only applies the empirical framework to a different national context, but also examines parental input using a different set of parenting measures. Each of these variations to the empirical application of the model provides insight into skill development and furthers our understanding of the role that parents play in this process.

Cross-Country Comparison: The first novel aspect of the empirical application presented in this chapter is that the model is estimated using Canadian data. In the previous chapter, I outlined why I chose to provide updated UK estimates by applying my model specification to an extended version of data that Hernández-Alava and Popli (2017) had already examined using a skill formation framework. For this chapter, it would have been possible to further prove my model's validity by applying it to US data in order to update the estimates provided by Cunha and Heckman (2008) and Cunha et al. (2010). While such updated estimates could provide valuable insight for skill development in the US context, there is more to be gained by estimating the model using a country that has not been previously examined using this type of methodology. Therefore, I have chosen to focus on Canada. This choice was motivated by both the desire to provide the first Canadian estimates for this type of model, and by the empirical value of providing a comparative analysis in a country that is known to have less income inequality than either the UK or the US.

Before exploring the specifics of the Canadian application of this model, I begin by explaining the factors which led me to identify Canada as a suitable context to provide international comparison for the UK analysis. The main motivating factor when selecting a second country for this thesis was to identify a country sufficiently similar to the UK and the US — so that the comparison is valid, but also sufficiently unique to yield contrasting results. From the existing literature of comparative studies, Canada and Australia were identified as suitable comparisons.[115]

In the broadest terms, the UK, the US, Australia, Canada and Britain are similar across a variety of economic, social and cultural factors. However, compared to the UK and the US, both Canada and Australia have lower levels of income inequality and

[115] These countries were chosen because they are English-speaking nations, identified as advanced economies, and have populations over 10 million.

this translates to smaller achievement gaps in childhood development. For example, compared to the UK and the US, Bradbury, Corak, Waldfogel, and Washbrook (2011) find that Canada and Australia both have smaller socio-economic gaps in cognitive and behavioural measures. By examining skill development in slightly different social and economic context, it is possible to understand what features of the model might be universal and which are driven by conditions specific to life in the UK — perhaps the role of parenting differs when socio-economic factors are less influential.

For this chapter, I focus on Canada; this is motivated by personal and practical considerations. As a Canadian researcher working in the UK, my perspective on education is shaped by both contexts, and I have a personal interest in understanding how my model applies to Canadian data. From a practical perspective, I am uniquely positioned to gain access to the data required for this study — the relevant data can only be analysed on site at specific Canadian universities, and international researchers are rarely granted permission to access it.[116] The comparative nature of this chapter meant that I was able to obtain support for this analysis from the Canada–UK foundation.

While the Canadian data explored in this analysis has been widely used in other research, to my knowledge, this is the first use of any dynamic model of skill formation in the Canadian context. As there exist no other Canadian studies using this method to measure skill development over the life course, my findings provide crucial evidence on the dynamic nature of skill development in Canadian children. These empirical insights are especially valuable for policymakers, as they allow for early childhood programs to be designed specifically for the Canadian context.

Examining Different Parental Inputs: In addition to examining skill formation within a different national context, this thesis also aims to see how the model can be applied to various measures of parental investment. Fortunately, the nature of the existing cohort studies means that each country chooses to focus on slightly different measures. As a result, each cohort study contains different measures of parenting, while still having sufficient overlap to measure the cognitive, non-cognitive and demographic features of the model. Due to the types of questions that are relevant to policy-makers around the globe, there is substantial overlap between the data contained in the MCS and that provided by cohort studies from other countries. Bradbury et al. (2011) compare cohort studies from Canada, Australia, the UK and the US to provide valuable insight into the comparability of such surveys. Their findings demonstrate the feasibility of using data from the Canadian-NLSCY and the UK-MCS in order to examine the same research question in two contexts.

[116] Statistics Canada limits data access to researchers who are either studying at a Canadian university, working for a Canadian organisation or have Canadian citizenship.

The Present Study

Based on the considerations discussed above, the present study provides three main contributions to the existing literature:

- Canada-specific estimates of the dynamic model of skill formation. While the skill formation model has been applied to UK and US samples, these international findings may not apply within the Canadian context.
- Use of existing parenting scales to capture different types of parenting behaviour. While the previous chapter presented the importance of differentiating between different types of parenting behaviour, the present study is able to use a pre-existing set of parenting measures to assess the effect of parenting on skill development. While a thorough exploratory factor analysis is done to confirm the structure of the parenting factors, the resulting parenting factors are in line with other research using the same data. This use of existing scales allows for contextualisation within other literature and validation of the modelling strategy.
- Evidence on the effect of specific parenting behaviours on skill development in Canadian children. Although the specific measures vary from the UK analysis, the results from the Canadian analysis contribute to the general understanding of skill development and inform our knowledge about the role of both family resources and parenting behaviours. Since both of these factors influence child outcomes, measuring them separately will allow me to identify the best areas for future public policy.

5.2 Data

The following section outlines the data used for the second empirical application of my PhD thesis. The section begins by discussing the choice of this specific data, followed by an overview of the survey. This description of the survey includes information about the survey sample, sample design and the strategies for data collection. I then discuss the selection of the specific subsample used in this analysis. The section concludes by reviewing the data within the context of the present study, including details about each of the measures used in the skill formation model, as well as summary statistics for the chosen variables.

Data Selection

As with the first empirical application, the Canadian analysis requires longitudinal data that: has an adequate sample size; includes data points from birth until adolescence; and contains sufficiently rich information on the measures required for a dynamic model of skill development. As discussed previously, *cohort studies* are the most suitable data for such an analysis. Chapter 3 provides more information about the use of cohort studies in education research.

In the Canadian context, the only national cohort study is the National Longitudinal Survey of Children and Youth (NLSCY). The NLSCY is a government-funded survey that is conducted through a collaboration between Statistics Canada and Human Resources and Skills Development Canada (HRSDC). Several provincial cohort studies exist, but these are focused on specific regions of Canada and their findings are less likely to be generalisable to the entire Canadian population. Furthermore, the sample size of the NLSCY is substantially larger than these small studies.

As well as being the only suitable Canadian option, the NLSCY meets all the needs of the present analysis. The data contained in the NLSCY is available to Canadian researchers, on application and has been used in a variety of educational research papers examining early childhood skill development and parental contributions in Canada (see for example: Baker, 2011; Bradbury et al., 2011; Lefebvre, Merrigan, & Verstraete, 2008; Waldfogel, 2007).[117]

[117]Details regarding access to the NLSCY data are provided later in this chapter.

National Longitudinal Survey of Children and Youth (NLSCY)

Survey Description

The National Longitudinal Survey of Children and Youth (NLSCY) aims to monitor the life experiences and development of a representative sample of Canadian children from birth through early adulthood. The NLSCY was designed to collect cross-sectional data measuring the biological, economic and social characteristics of Canadian children; and to collect longitudinal data that can be used to assess the impact of various factors on a child's development. Both the cross-sectional and longitudinal data are intended to offer accurate information to policy-makers and other groups and individuals that are involved in promoting the well-being of Canadian children (Statistics Canada, 1996).

Data contained in the NLSCY was collected using a variety of methods. Figure 5.1 depicts the types of data collection that Statistics Canada used during the NLSCY. This diagram highlights the complexity of the data collection as well as the multiple sources involved in each stage of the study. More detail is provided in the NLSCY User Guides, as well as later in this section.

Figure illustrating the survey design of the NLSCY removed for copyright reasons.

Copyright holder is Statistics Canada.

For original figure, refer to Michaud (2001)

Fig. 5.1 NLSCY Survey Design – diagram obtained from p. 400 of Michaud (2001)

The first round of data collection for the NLSCY took place in 1994/1995, when the children in the sample were 0 to 11 years old. The sample was revisited and data was collected every two years. Statistics Canada refers to each period of data collection as a *cycle*. The NLSCY consists of eight cycles.

The NLSCY required substantive survey design, ethical approval and pilot-testing, and contains thousands of variables for each cycle of the survey. Full information about the available measures, the structure of the survey and other factors can be obtained from Statistics Canada but are beyond the scope of this PhD thesis. This section discusses the procedure required to access the data, provides a brief discussion of the selection of the NLSCY sample and my selection of the applicable subsample and concludes with an overview of the variables included in the present analysis.

Data Access

Data managed by Statistics Canada is subject to the *Statistics Act* and is only accessible under strict conditions.[118] To ensure all conditions of data access are met, Statistics Canada limits use of confidential data to secure locations, known as Research Data Centres (RDC). Data held in the RDCs can only be accessed in person, and is subject to stringent regulations to protect respondents' confidentiality. These RDCs are housed in Canadian universities and impose strict regulations on data usage.

To access the data I spent a total of six months conducting analysis at the Prairie RDC, which is located in Calgary, Alberta, Canada. As a visiting student, I obtained support from a professor at the University of Calgary who endorsed the application that I submitted to the Social Sciences and Humanities Research Council (SSHRC).[119] Once approval for my project was granted, I was required to complete police background checks and become a 'deemed employee' of Statistics Canada before finally gaining access to the data.

The micro-data is held on secure servers within the RDC computer lab. These computers have no access to the internet or outside networks. All analysis of the data must take place within the RDC, and no information may leave the RDC without being vetted by the research analysts that oversee each location. The vetting process takes 6 to 8 weeks and is designed to protect confidentiality of respondents. The regulations for vetting and data release are extensive and generally not pertinent to the reader.

Several regulations are worth noting as they prevent the presentation of results in their usual format; these regulations concern *cell count* and *weighting*. Regarding cell count, the rules dictate that no results may identify a subgroup of five or fewer individuals; this means that some summary statistics cannot be reported, and that certain

[118] Legally, certain confidential data can only be accessed for research projects led by a Canadian citizen or permanent resident, and all researchers must meet the requirements for security clearance.

[119] Appendix B.2 provides a copy of this application.

covariances are omitted from results tables. Similarly, histograms and other descriptive statistics must be created in such a way that this requirement is met. Additionally, RDC regulations only permit the release of weighted results. Thus, any results presented in this chapter are calculated using the survey weights provided by Statistics Canada.[120]

Ethical Considerations

As with the first empirical application, it is important to acknowledge that the present study involved human participants, even though I did not interact with them. Below is an overview of the ethical considerations taken by Statistics Canada concerning consent and confidentiality, as well as the ethical considerations of my own analysis.

A major ethical concern with any survey is that consent is obtained prior to data collection. Statistics Canada operates under the assumption that implicit consent is given by agreeing to participate in a voluntary survey (Statistics Canada, 2010). Respondents are reminded of the voluntary nature of the NLSCY at the beginning of each cycle, as well as before certain sensitive questions. Furthermore, respondents are provided information about the intended use of their data and that their identity will remain confidential. With this information, implicit consent is assumed for all participants in the NLSCY. As young children are not capable of directly giving consent, this meant that informed consent was obtained from the primary caregiver of the child. At each cycle, additional written consent was required in order to collect information from the child's teacher and school. For direct child assessments, further verbal parental consent was obtained at the time of the assessment. In addition to the parental consent obtained by the Statistics Canada, verbal assent was obtained from the child for any direct assessment.

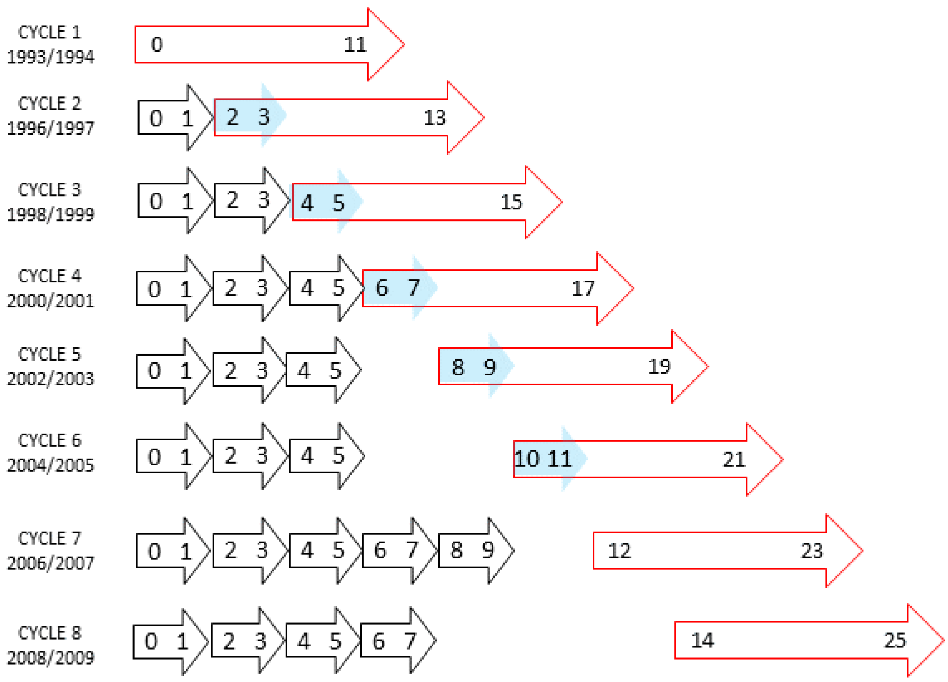
The other major ethical concern for the NLSCY is the confidentiality of survey respondents. As mentioned above, the micro-data provided by Statistics Canada in the Research Data Centres (RDC) is subject to multiple security measures in order to preserve the confidentiality of individual respondents. Within the RDC, the NLSCY is stripped of certain identifying information and instead identifies each participant using an ID number. Not only does this ID number allow for data linkage between the survey cycle, it also limits the ability to identify individual respondents.

Furthermore, any results that are released from the RDC must meet the confidentiality requirements set forth by Statistics Canada. This includes the use of survey weights to adjust for population characteristics and not releasing results that describe small demographic subgroups. In addition to the requirements laid out by Statistics Canada, I have limited discussion which focuses on specific subgroups. Though controls are included in the analysis, specific coefficients are only reported when necessary.

[120] More information about these specific weights is provided later in this chapter, as well as general information about the use of survey weights provided in Chapter 3.

Survey Sample

The target population for the NLSCY was all non-institutionalised children residing in the ten Canadian provinces[121], but excluding those living on native reserves or whose parents were full-time members of the Canadian Armed Forces. The NLSCY includes several different samples that were each targeted for a specific research purpose. The 'original cohort' refers to the children initially sampled in 1994. The original sample for the NLSCY was identified as a subsample of the 1994 Canadian Labour Force Survey (LFS) sample.[122] The 'early childhood development' (ECD) cohorts were selected during later cycles and only followed for three or four survey cycles. The present study focuses on the original cohort as it requires data that follows children for at least six cycles of data collection. This sample structure is illustrated in Figure 5.2. In addition to the original cohort (long arrows), the later cycles include smaller arrows representing each of the ECDs.[123]



Notes: Adapted from "NLSCY Survey Overview for the 2008/2009 Data Collection – Cycle 8" (p. 13)
Each arrow represents the age range (in years) of a subsample of the NLSCY cohorts
Red arrows represent the original NLSCY cohort
Blue shading represents the sub-sample used in this thesis
Shorter (black) arrows represent the Early Childhood Development Cohorts (ECD)

Fig. 5.2 NLSCY Survey Structure

[121] Canada has ten provinces and three territories. The three territories are often excluded from national data as their small population and remote location makes it difficult to collect data and maintain anonymity.

[122] The LFS is a monthly survey used to provide standard employment measures within Canada.

[123] The later cycles of the NLSCY included new samples of children aged 0–1 at the time of each cycle and followed them for the next two cycles. The aim of surveying these additional children was to monitor the early years of Canadian children, and they are collectively referred to as the Early Childhood Development (ECD) cohort. While children in the ECD cohort were only contacted for three cycles, the original sample was surveyed for all eight of the NLSCY cycles.

The original cycle of the NLSCY, Cycle 1, included information about 22,831 children who were aged 0–11 in 1994. The data included a maximum of four children per household surveyed. In Cycle 2, Statistics Canada decided to reduce the size of the longitudinal cohort due to budgetary constraints. To reduce the sample, Statistics Canada limited collection to a maximum of two children per household surveyed as well as excluding Cycle 1 respondents who were also a part of the National Population Health Survey (NPHS). This reduced the Cycle 1 cohort to the longitudinal cohort of 16,903 children, which corresponds with a response rate of 86.7% of the redefined target sample.

For Cycles 2–8, the NLSCY attempted to collect follow-up information for these 16,903 children. Families who had not responded for two consecutive cycles were no longer contacted. The follow-up surveys ended with Cycle 8 in 2008–2009, which included responses from 10,208 of the children originally surveyed in 1994.

The present analysis focuses on Cycles 3–6 the NLSCY. Table 5.1 shows the number of respondents to each of these first six cycles as well as the age of the cohort at each period. Cycle 6 contains information on 11,178 of the children originally surveyed in 1994. The exact size of the NLSCY sample over Cycles 1–6 is shown in Table 5.1.

Table 5.1 Number of Respondents: NLSCY Longitudinal Cohort

Year	NLSCY Cycle	Full Longitudinal Cohort				Relevant Longitudinal Cohort		
		Child Age (in years)	Number of Respondents[1]	Response Rate (%)	% of C1 [2]	Child Age (in years)	Number of Respondents[3]	% of C1[2]
1994/5	C1	0–11	16,903	86.7	–	0 – 1	4,052	-
1996/7	C2	2–13	15,391	79.1	91.5	2 – 3	3,740	92.3
1998/9	C3	4–15	14,777	76.0	88.7	4 – 5	3,638	89.8
2000/1	C4	6–17	13,176	67.8	78.7	6 – 7	3,229	79.7
2002/3	C5	8–19	12,280	63.1	74.1	8 – 9	3,157	77.9
2004/5	C6	10–21	11,178	57.6	67.9	10 – 11	2,825	69.7

Notes:

[1] C1 had 22,831 respondents. Due to budgetary constraints the sample size was reduced to 16,903 at the start of C2.

[2] No detailed information is available on the age of children in the original target sample. Thus, the value of "% of Cycle 1" is calculated to show attrition.

[3] For Cycles 1–5, the number of respondents in the relevant age-range includes children who have left the country or died since Cycle 1. This number is small (79 children as of Cycle 5).

Source: Statistics Canada — NLSCY User Guides

As well as the standard response rate, Table 5.1 includes a measurement for the current respondents as percentage of Cycle 1. Since the target sample was not based on specific age-ranges within the larger target of ages 0–11, this measure allows for the comparison of specific age-ranges to the full sample. It can be seen that the relevant longitudinal cohort follows a similar pattern of response to the full longitudinal sample, with 2,825 children who were ages 0–1 in 1994 responding to Cycle 6 of the NLSCY.

Sample Design

Though the sampling strategy used by Statistics Canada to select the NLSCY sample was beyond my control, an understanding of the methods used is crucial to any analysis dealing with the NLSCY data. As discussed in the previous empirical application, providing a set of representative results is one of the primary aims of my empirical studies, and the sample design is what determines the representativeness of a sample. In Chapter 3, I have shown how representative samples need to be designed in such a way that they are not only representative of the larger population, but also contain large enough subsamples of various subgroups within the population. Below is a brief explanation of the sample design used to select the NLSCY sample and the impact it has on the present analysis. More detailed information on the NLSCY sampling strategy can be found in Statistics Canada reference material.

Like the MCS, the NLSCY used a clustered and stratified sample design. This meant the NLSCY oversampled some regions and demographic subgroups in order to obtain a large enough sample of small subpopulations in Canada. More specifically, HRSDC and Statistics Canada aimed to have "sufficient sample at the national level to reliably measure characteristics with a national prevalence of 4% for each age group after five cycles" (Michaud, 2001, p. 405). An explanation of this stratification and clustering is provided in Chapter 3. Further details regarding the NLSCY's specific application of these sampling methods are provided in Appendix B.

The complex survey design of the NLSCY means that there are several considerations that must be made when using the data that it contains. First, the distribution of characteristics observed in the NLSCY data may not reflect the actual distribution of characteristics in the Canadian population as certain groups are purposefully oversampled. Second, the majority of statistical techniques depend upon the assumption that a sample is randomly selected, and the NLSCY sampling design violates this assumption. Fortunately, there are statistical techniques that can be used to make any findings more representative of the target population and to satisfy the assumption of randomly sampled data — these will be discussed in greater detail within my analysis.

As with all longitudinal survey data, the NLSCY is subject to attrition and item-non-response. The overall response rate of the NLSCY is presented in Table 5.1, but item-non-response means that the percentage with complete responses to all cycles is even lower than the values presented. The overall response rate for the NLSCY is slightly lower than that of the MCS, but still well above the response rates of 20–30% generally seen in longitudinal surveys of this size (Cohen et al., 2007).

Survey attrition and item-non-response further magnify the fact that the sampling strategy of the NLSCY may mean that traditional estimates of population parameters do not reflect the reference population. To make the data suitable for representative analysis, the NLSCY data includes final sampling weights that intend to account for sampling strategy and attrition. For further information on the use of sampling weights, the reader should refer to Chapter 3.

Included in the NLSCY data for each cycle are standard sampling weights and bootstrap weights. Statistics Canada dedicates substantial resources to ensure that the NLSCY survey weights represent the known number of children by age, gender and province. These survey weights are determined by the original probability of being selected and are then readjusted at each cycle to account for survey non-response and to match the most recent counts of Canadian children.

The sampling weights are designed to provide estimates that mirror the reference population. For example, rural children are overrepresented in the NLSCY sample, and therefore they would be assigned a lower sampling weight, while urban children are under-sampled and are thus assigned higher sampling weights. A weighted analysis accounts for the fact that observations with the higher sample weights need to be counted more times than those with a lower sample weight. Using sampling weights provides point estimates for the population that reflect the true composition of the population, whereas a simple point estimate would put too much emphasis on the traits of rural individuals (as they are overrepresented) and neglect the responses of urban individuals (who are under-represented).

While traditional survey weights are able to provide adjusted point estimates, they do not sufficiently correct for sampling variance. The NLSCY's complex sample design, non-response adjustments and post-stratification make it impossible to calculate the sampling variance using traditional econometric methods. Instead variance estimates can be calculated with the use of bootstrap weights. In addition to the sampling weights discussed above, Statistics Canada also provides a set of 1,000 bootstrap weights for each respondent. Re-estimating the model of interest separately using each of these bootstrap weights provides 1,000 estimates for each parameter. The variance of these 1,000 estimates provides a reliable estimate of the sampling variance for the estimate obtained using the original sample weight.

Further information on the NLSCY sample design can be found in the user guides that accompany the data. To account for the complex sample strategy, as well as attrition from the first cycle to the last, sampling weights and bootstrap weights are provided by Statistics Canada. To comply with confidentiality requirements imposed by Statistics Canada, all results reported in this thesis make use of the provided sample weights. The NLSCY's use of sample weights is discussed in greater detail in Appendix B.4, with the use of bootstrap weights presented in Appendix B.5.

Data Collection

Once the NLSCY target children were identified using the LFS, data was collected directly from the survey respondents. Earlier in this chapter, Figure 5.1 provides a visual representation of how data was collected for each cycle of the NLSCY. For each cycle, data was collected from multiple sources for each child.

Unlike the MCS, which collected data from the child's mother, the NLSCY identified the person most knowledgeable (PMK) about the child at each time of sampling. The data included in the NLSCY is based on a series of questions asked of the PMK, and where applicable, a second series of questions asked of the spouse of the PMK. This data was collected face-to-face with the interviewer using computer-assisted interviewing (CAI). Each cycle of the NLSCY also included direct cognitive assessments of the child. Once the child reached age 10, they were given self-complete questionnaires in their homes. Questionnaires were also sent to the child's teacher and school principal.

Selection of Relevant Subsample

Since the model used in this analysis focuses on development from early childhood through adolescence, it only requires the first six cycles of the NLSCY data. Respondents must have responded to each cycle, as the model cannot be applied to individuals missing an entire time point in the model. Additional sample restrictions were required to provide a sample that met the requirements of the present analysis. This subsample was designed to maximise the amount of available information within the constraints of the available data. After meeting all the requirements of the model, the final usable sample for the present analysis was 1,234 children. This sample was restricted to include:

- children who had completed Cycles 1–6 (9,854 households).
- children who were age 0–1 at time of C1 (2,510 cases).
- singleton children: no twins or triplets (2,368 cases).
- those whose respondent for every cycle is a parent (2,276 cases).[124]
- children with full responses to the relevant demographic questions (2,258 cases).
- children who had responses to the cognitive assessments in Cycles 2–6 (1,308).
- children who had responses to at least two of the behaviour measures and to at least two parenting/investment questions (1,234).

Due to the confidentiality requirements of Statistics Canada, the same identical sample was maintained for all analyses.[125]

[124] With the exception of 92 cases, all respondents were the mother or father.

[125] The pattern of partial response in the NLSCY meant that if certain measures were not included in a particular model, the sample size could have been increased by 10–20 respondents. Properly reporting the results from such slight changes in sample size would have required provision of descriptive statistics for each relevant subsample, and would consequently violate Statistics Canada regulations.

Measures Used in Present Analysis

Demographic Measures

As with the first empirical application, this analysis takes care to avoid attributing to parental investment what should be attributed to other factors correlated with cognitive and non-cognitive ability. Thus, the analysis must control for: differences in the initial skills of children; family characteristics that may influence development; and systemic differences in the measurement of the observed variables in different demographic subgroups. The reasons underpinning these controls are discussed in greater detail in the previous two chapters.

While the types of controls are the same across different empirical applications of the model, the specific choice of demographic controls depends on the context in which the data was collected. For example, in the UK analysis, it was assumed that written test scores of non-English speakers are likely to systematically underestimate these children's true ability, so an analysis must control for English language ability. In Canada, there are two official languages, and tests are administered in the preferred of these two languages, so instead the analysis must control for individuals who speak neither French nor English at home. Only controlling for English ability would include children who are from French-speaking households along with those who were from foreign language households — and thus would fail to control for the measurement error of language.

To control for the sources of bias discussed above, the Canadian analysis includes a very similar list of demographic measures to the UK analysis, but with slight modifications to match the Canadian context and the data available in the NLSCY. Chapter 4 has already discussed the reasoning behind the demographic measures that are consistent with the UK application, and for brevity these explanations are not repeated in this chapter. Instead, I focus here on the changes made to the demographic variables in order to make them applicable to this Canadian sample.

The first change was the inclusion of a binary variable to indicate the gender of the individual completing the survey, as the PMK was not necessarily the child's mother. The second change is the inclusion of a binary variable indicating whether either of the child's parents was born outside Canada. This controls for some of the differences that might be seen in immigrant families beyond the impact of language and ethnicity. Finally, the language variable is modified as described above to account for Canada's bilingual use of French and English. All this results in the control variables listed in Table 5.2.

Table 5.2 Definitions of Included Covariates

Variable	Description
Time Invariant [1]	
Child's Gender	Binary variable set to equal one if the child is female.
Birth Weight	Child's weight at birth, reported by the PMK.
Mother's Age at Birth	Biological mother's age at the birth of the child.
Non-Native Speaker	Binary variable set equal to one if language spoken in the home is "neither English nor French".
Either Parent Overseas-Born	Binary variable set equal to one if either parent was born outside Canada.
Gender PMK	Survey was completed by Person Most Knowledgeable (PMK) about the child, who may not be the mother. This indicates gender of this respondent.
Child's Ethnicity	PMK asked to define child's ethnicity. Set equal to one for those who chose 'white.'
PMK's Highest Educational Qualification	The educational attainment of the PMK as of 1994. Defined using six categories.
Spouse of PMK's Highest Educational Qualification	The educational attainment of the PMK as of 1994. Defined using six categories.
Time Variant	
Single Parent Household	Binary variable, set equal to one if PMK reported 'Child lives with - one parent only'.
Number of Siblings	As reported by the PMK.
Household Income Quintile	Categorical variable, which places the family into one of five income quintiles based on their income in the previous calendar year. Income is calculated using consumer price index (CPI), number of people in the household and household income.

Notes:

[1] With the exception of ethnicity, all 'time invariant' variables were determined using Cycle 1 data. Ethnicity was not asked until Cycle 2, so this response was used.

While Table 5.2 introduces the relevant covariates, the present section outlines exactly how they were measured by Statistics Canada and the way in which they were coded in the present analysis. Like the MCS, the NLSCY measured demographic variables during all cycles of the survey, but some measures were only collected during the initial period of data collection. Responses from Cycle 1 were used for these time-invariant variables for all portions of the analysis. These time-invariant measures are defined as:

- **Child's Gender:** Equal to one for female and zero for male.
- **Birth Weight:** The PMK was asked to report the child's weight at birth. An additional indicator variable was created to indicate the child was low-birth-weight — this was defined as being less than 2,500 grams at birth. Using this binary indicator for analysis is beneficial as this is a recognised level at which health problems are more likely and a continuous measure would mean that the difference between 2,400

grams and 2,500 grams is treated the same as the difference between 2,800 and 2,900.

- **Biological Mother's Age at Birth:** Unlike many of the other NLSCY variables which refer to the PMK, this specifically refers to the biological mother. Some respondents were missing this information for Cycle 1. For these respondents, the variable was based on the responses from later cycles if available.
- **Non-Native Speaker**: This is slightly different to the language variable included in the MCS. The respondent was asked "the language first learned in the home and still understood." If their response was neither French nor English, they were categorised as a non-native speaker, and assigned a value of 1 for this variable.
- **Either Parent Overseas-Born:** Nearly 25% of Canadians are foreign-born, and these individuals may have different experiences with the Canadian education system and respond differently to the parenting measures described in this analysis. This control variable is based on the place of birth reported by the PMK. Unlike measures of ethnicity or citizenship, this measure directly reflects the individual's background and exposure to Canadian culture. In a multicultural country where citizenship is available to many foreign-born residents, this is an important distinction.
- **Gender of PMK:** This variable helps define the proportion of respondents who are the children's mothers, and those who are the children's fathers, because the subsample was limited to only those who had a parent as a primary respondent, this can be interpreted as the portion of respondents who are the child's mother.
- **Child's Ethnicity:** The PMK was asked, "How would you describe [your child's] race or colour?". They could choose more than one category. In the micro-data file, there are multiple response categories for ethnicity. Due to confidentiality concerns, these summary statistics use binary indicator equal to one if the child is white.
- **PMK's Highest Educational Qualification:** The interviewer asked the PMK to advise the highest level of education they had attained. The response was then allocated to one of 9 categories. For the present analysis, these categories were collapsed into six categories in order to preserve the confidentiality of those in some of the smaller categories. Each of the six categories are coded using binary variables equal to one for the level completed and zero otherwise. The categories are:
  - *Higher Degree*
  - *Bachelor's Degree*
  - *Trade School or Community College Diploma*
  - *Some Post-Secondary*
  - *High School Graduate*
  - *Less than Highschool*
- **Spouse of PMK's Highest Educational Qualification:** Given that a child is influenced by both parents, the spouse's highest level of education was also included. The categories used are the same as those used for the PMK.

Table 5.3 provides the summary statistics for the demographic characteristics of the sample used in the analysis as compared with all of the NLSCY respondents who were in the relevant age range.[126] For the binary variables (e.g. ethnicity, child's gender), the mean represents the percentage of the sample possessing the measured feature.

Table 5.3 Time Fixed Demographic Measures: Weighted Descriptive Statistics

	Analysis Sample[1]		Full NLSCY Sample[2]	
	mean	SD	mean	SD
Child's Gender (=1 if female)	0.507	0.500	0.485	0.500
Birth Weight (kilograms)	3.423	0.591	3.415	5.989
Low Birth Weight (<=2,500 grams)	0.059	0.236	0.066	0.248
Mother's Age at Birth	29.252	4.517	29.482	7.239
Non Native Speaker (French or English)[3]	0.099	0.299	0.111	0.314
Either Parent Overseas-Born	0.221	0.415	0.233	0.423
Gender of PMK (=1 if female)	0.920	0.272	0.918	0.274
Child's Ethnicity (=1 if white)[4]	0.899	0.302	0.875	0.423
Highest Educational Qualification (PMK)				
Less than Highschool	0.133	0.339	0.158	0.365
Highschool Graduate	0.159	0.366	0.152	0.359
Some Post-Secondary	0.279	0.449	0.275	0.447
Trade School/Community College Diploma	0.216	0.412	0.225	0.418
Bachelor's Degree	0.178	0.383	0.154	0.361
Higher Degree	0.035	0.185	0.035	0.183
Highest Educational Qualification (Spouse of PMK — if applicable)				
Less than Highschool	0.147	0.354	0.174	0.379
Highschool Graduate	0.189	0.392	0.174	0.379
Some Post-Secondary	0.225	0.418	0.234	0.423
Trade School/Community College Diploma	0.233	0.423	0.223	0.416
Bachelor's Degree	0.157	0.364	0.147	0.354
Higher Degree	0.049	0.215	0.049	0.216
Observations	1,234		2,510	

Notes:
[1] Children age 0–1 during Cycle 1 who: completed Cycle 1–6, were singleton children, respondent for all cycles was a parent, had data for demographic, cognitive, behaviour and parenting measures.
[2] Children age 0-1 at the time of Cycle 1 who: completed Cycle 1–6.
[3] Canada is officially bilingual, so this variable refers to families who speak neither English or French as the primary language in the home.
[4] Due to confidentiality requirements, it is not possible to provide more specific categories of ethnicity.
[–] All descriptive statistics are calculated using provided survey weights and bootstrap weights.

When comparing the two groups, it can be seen that there are slight differences in the composition of analysis sample compared to the full NLSCY sample. Specifically, the analysis sample has a higher proportion of bachelor's degrees, and around 1% fewer non-native speakers, overseas-born parents and non-white children. These similarities in sample composition are likely due in part to the use of sampling weights, which are designed to adjust for the demographics of the national population.

[126] RDC regulations only allow weighted sample statistics to leave the data centre. As such, it is not possible to present the unweighted descriptive statistics for the analysis sample. However, unweighted descriptive statistics were examined in the RDC and these do not vary notably from the results presented.

The covariates described above remain fixed over a child's lifetime, but other covariates relevant to the analysis change over time and must be collected for each stage of development included in the analysis. These time-variant measures were collected at each of the NLSCY cycles and are defined as:

- **Single Parent Household:** The PMK was asked with whom the child lived with. This was used to define the child's 'single parent status' — meaning whether the child lived with 'two parents' or 'one parent only'. This was used to create a binary variable equal to one if the child lived with 'one only'.
- **Number of Siblings:** The PMK was asked to report how many siblings lived in the house, including full, half, adopted and foster.
- **Family Income Quintile:** Income quintiles based on equivalised real household income. The NLSCY included the total household income. I adjusted this value using consumer price index (CPI) and the reported number of people in the household. This value was used to calculate weighted income quintiles for the original NLSCY sample which were then included and carried forward to the final analysis sample.

Table 5.4 provides the weighted descriptive statistics for the time-varying covariates relevant to this analysis. As with the MCS, the number of siblings rises as the child ages. This simply indicates that families are having other children after the cohort child. Again, the rate of single parent families increases slightly as the child ages — a pattern that is in line with general demographic trends. Even after using the survey weights, the NLSCY sample still shows under-representation of the lowest income quintile.

Table 5.4 Time Varying Demographic Measures: Weighted Descriptive Statistics

	Cycle 1 *Age 0/1*		Cycle 2 *Age 2/3*		Cycle 3 *Age 4/5*		Cycle 4 *Age 6/7*		Cycle 5 *Age 8/9*	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Age in Months	11.879	6.651	35.521	6.751	58.001	6.655	84.473	7.132	104.931	6.784
Number of Siblings	0.893	1.000	1.147	0.958	1.300	0.884	1.416	0.947	1.434	0.925
Single Parent	0.061	0.239	0.062	0.242	0.093	0.290	0.107	0.309	0.122	0.328
Annual Income($)[1]	39,531	25,046	38,497	25,342	43,344	32,166	46,339	31,290	48,618	35,018
Household Income[2]										
Lowest Quintile	0.147	0.354	0.145	0.352	0.167	0.373	0.191	0.393	0.162	0.369
Second Quintile	0.186	0.389	0.217	0.412	0.183	0.386	0.188	0.391	0.250	0.433
Third Quintile	0.210	0.407	0.211	0.408	0.220	0.414	0.232	0.422	0.191	0.393
Fourth Quintile	0.232	0.422	0.221	0.415	0.221	0.415	0.171	0.377	0.209	0.407
Highest Quintile	0.226	0.419	0.206	0.405	0.209	0.407	0.218	0.413	0.188	0.391
Observations	1,234		1,234		1,234		1,234		1,234	

Notes:

All descriptive statistics are calculated using provided survey weights and bootstrap weights.

[1] Equivalised Real Household Income, calculated using CPI (constants), number of people in household, and reported household income.

[2] Income quintile calculated as quintile ranking compared to all respondents for given NLSCY cycle.

Cognitive Measures

At each cycle, the NLSCY conducted a variety of cognitive assessments. These were administered directly to the child by the interviewer. The exact assessment used was dependent on the age of the child at the time of the survey. Table 5.5 lists the available cognitive measures, for the relevant age group, from Cycles 2–6 of the NLSCY. For Cycle 4, two cognitive measures are listed; indicating that the relevant subsample falls on the age boundary for the cognitive tests and that the younger part of the sample has been given one test, while the older group has been given another.

Table 5.5 Cognitive Measures Included in the NLSCY

	NLSCY Cycle				
	C2 *Age 2/3*	C3 *Age 4/5*	C4 *Age 6/7*	C5 *Age 8/9*	C6 *Age 10/11*
Motor and Social Development (MSD)	X				
Peabody Picture Vocabulary Test (PPVT)		X	X		
Canadian Achievement Test (CAT)			X	X	X

Note:
[1] During Cycle 4, 6 year old children were given the PPVT, while the 7 year old children were given the CAT.

Further details about these cognitive measures are provided below. While it would be ideal to have the same measures repeated over time, this is not possible with the measures included in NLSCY. Fortunately, NLSCY cognitive measures are independently validated and have been used in much research.

Motor and Social Development (MSD): Children aged 0 to 3 were assessed using the Motor and Social Development (MSD) scale. This test was originally created for the National Health Interview Survey in the US and is designed to measure motor, social and cognitive development in children from birth to 3 years of age[127] (Center for Human Resource Research, 2000). The MSD was also included in the National Longitudinal Survey of Youth (NLSY), an American cohort study, and was one of the cognitive measures used by Cunha et al. (2010) in their estimation of the skill formation model. Though they use it in their analysis, Cunha et al. (2010) note that the MSD is a noisy measure, especially when compared with other cognitive measures [p. 909].

The MSD scale is created using a set of 15 'yes' or 'no' questions asked of the parent about whether a child has ever performed a specific behaviour. These questions are selected from 48 motor and social development items, with the specific questions asked being contingent on the child's age. The MSD score is the sum of the number of 'yes' answers to the 15 questions. In the NLSCY, the MSD questions were included in the PMK questionnaire. Since the specific questions asked varied depending on the

[127] Though the test was targeted at children aged 0 to 3, it did include measures for children under age 5 (up to, and including, 47 months of age).

age of the children, Statistics Canada found that children in different age-ranges had different average scores.[128] The strong influence of age in the NLSCY matches the findings in the American NLSY. Researchers are told that "caution should be exercised when interpreting the [MSD] results" as "the distribution of scores for NLSY79 children on this assessment tends to be more peaked for the youngest and oldest children" (Center for Human Resource Research, 2000, pp 45-46).

Peabody Picture Vocabulary Test (PPVT): The Peabody Picture Vocabulary Test — Revised (PPVT–R) was introduced by Dunn and Dunn (1981) to measure receptive and expressive vocabulary in individuals aged 3 and older. The NLSCY included a French language adaption of the PPVT-R known as the *Échelle de vocabulaire en images de Peabody* (EVIP), allowing children to be assessed in either official language.

For the NLSCY, the PPVT-R was administered to children ages 4 and older who had not yet reached grade two. The child was presented with four black and white images and asked to identify the one which matched the word spoken by the interviewer. The raw score corresponds with the total number of correctly identified images. This raw score allows researchers to compare a child's progress over time. It is assumed that an older child is able to identify more images correctly.

Standardised scores are also included in the NLSCY data. The standardised PPVT–R scores were calculated using the NLSCY sample as the norming population and assigning standard scores so that the scores had a mean of 100 and standard deviation of 15 for each one-month age group. As the final analysis will include age controls, the raw scores were included in the model for this paper. The standardised scores are presented below in order to show how the subsample compares to the larger NLSCY sample.

Canadian Achievement Tests (CAT): The final cognitive measure used in the analysis is the Mathematics Computation Exercise, which is an abridged version of the Mathematics Computation Test included in the Canadian Achievement Tests, Second Edition (CAT-2). As the test included in the NLSCY is taken from the CAT-2, it is often referred to as the CAT. This mathematics test was designed to measure a child's grade-specific mathematics skills, including measures of addition, subtraction, multiplication, division, decimals, fractions, negatives and exponents. Within the NLSCY, the CAT consisted of 20 questions that corresponded with the grade level of each child tested. The child's raw score is simply the number of correct answers.

The CAT results are also reported as a classical scaled score and item-reponse-theory (IRT) scaled score. The present analysis uses the classical scaled scores as it is based on a normative sample created to be reflective of the Canadian population. Use of this score allows for comparisons outside the NLSCY sample, as opposed to the IRT score which

[128] Age-standardised scores were created with a mean of 100 and standard deviation of 15 for each one month age group (Statistics Canada, 2004).

is only a relative ranking within the NLSCY sample. Additionally, Statistics Canada stopped providing IRT scaled scores in Cycle 7 of the NLSCY. By using the classical scaled scores this paper allows for comparability with later cycles.

The classical scaled score is derived using norming tables provided by the Canadian Testing Centre (Statistics Canada, 2004) to adjust for the child's current grade in school. These norming tables are based on a representative sample of English-speaking Canadian children in 1992 who took the extended version of the Canadian Achievement Tests. The scaled scores use a single scale for all grade levels, allowing researchers to see a progression in the child's score over time. The reference material included with the NLSCY advocates the use of either of the scaled scores, but explains that "rank test analysis performed using both methods of scoring showed no significant difference between the two measurements." (Statistics Canada, 2004, p.148)

Cognitive Scores in the NLSCY Sample

Table 5.6 provides summary statistics for the raw and adjusted cognitive scores included in the NLSCY. My analysis uses the raw scores for the MSD and PPVT; and the classical scaled scores for the CAT. To allow for the use of different scores within the same cycle, I convert the NLSCY provided cognitive scores to z-scores before including them in the analysis. The NLSCY adjusted results are presented in Table 5.6 as they give a clearer sense of how the NLSCY sample performs on the various measures.

Table 5.6 Measures of Cognitive Ability: Weighted Descriptive Statistics

	Raw Scores		Scores Standardised by NLSCY		
	Mean	SD	Mean	SD	Obs
Cycle 2 (Age 2/3)					
Motor and Social Development (MSD)[1]	11.538	2.829	101.160	14.749	1234
Cycle 3 (Age 4/5)					
Peabody Picture Vocabulary Test (PPVT)[2]	57.832	18.088	99.983	13.922	1234
Cycle 4 (Age 6/7)					
Peabody Picture Vocabulary Test (PPVT)[2]	84.918	16.096	102.543	12.533	708
Canadian Achievement Test (CAT)[3]	N/A[4]	N/A[4]	293.710	38.302	526
Cycle 5 (Age 8/9)					
Canadian Achievement Test (CAT)[3]	N/A[4]	N/A[4]	363.946	48.452	1234
Cycle 6 (Age 10/11)					
Canadian Achievement Test (CAT)[3]	N/A[4]	N/A[4]	421.900	53.541	1234

Notes:
[1] MSD: raw scores represent number of correct answers and range from 0–15. Standardised scores are adjusted so that each one-month age group has a mean of 100 and SD of 15.
[2] PPVT: raw scores represent number of correctly identified images. Standardised scores are adjusted so that each one-month age group has a mean of 100 and SD of 15.
[3] Reported scores are classical scaled scores, for further information see the NLSCY documentation provided by Statistics Canada.
[4] For the Canadian Achievement Tests, the questions in the test were adjusted based on the child's grade level, so the raw score alone provides no relevant information.
[5] All reported scores calculated with sample weights and bootstrapped standard errors.

For the MSD, the NLSCY population has a sample mean of 101.160 which falls slightly above the reference population's mean of 100. As the norming population was the larger NLSCY sample, the subsample used in this analysis performs slightly higher than the full sample. In Cycle 3, the PPVT scores closely align with the reference mean, while at Cycle 4 the sample mean of 102.543 is above the reference population. Though they do not perfectly align with the reference ranges, these scores show that, using the provided sampling weights, the subsample's performance is roughly what would be expected. This finding is important, as there is always potential bias which arises from conducting an analysis on a selected subsample of the data. Specifically, since the provided sampling weights are based on the full NLSCY sample, they may not adequately adjust for the subsample which I have selected. Finally, for the CAT scores, the classical scaled scores rise with age, which is to be expected as this scale is designed to capture a child's progress over time.

Histograms of the relevant variables are shown in Figure 5.3. These histograms exclude extreme outliers in the tails in order to comply with confidentiality requirements. Due to Statistics Canada regulations, it is not possible to release histograms for both age-standardised and raw scores. The histograms for the specific scores used in the present analysis are presented in Figure 5.3, because they provide more relevant information on the outcomes.

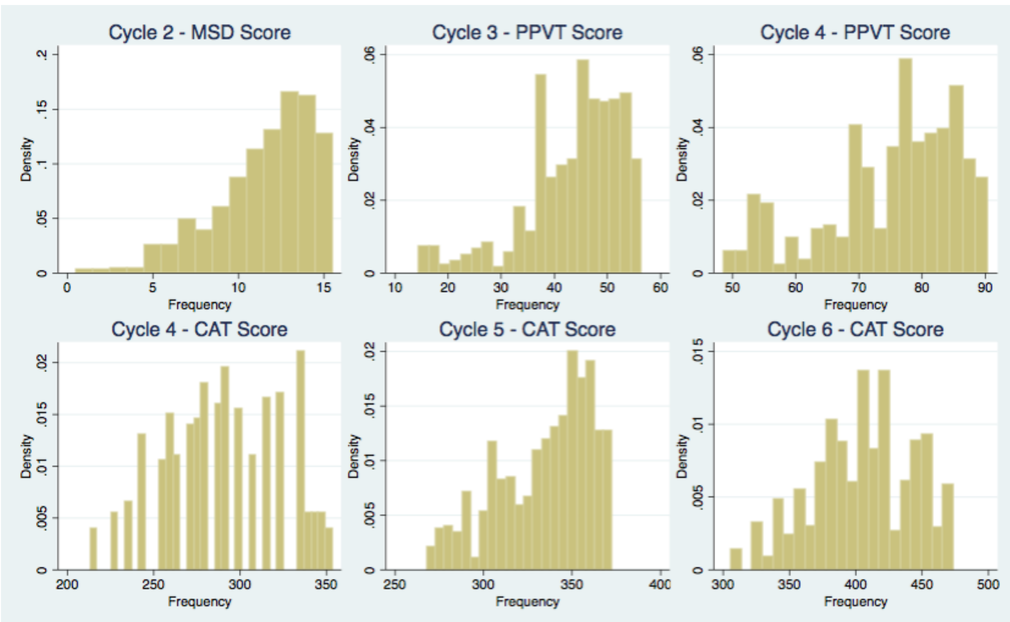


Fig. 5.3 Distribution of Cognitive Scores: NLSCY Cycles 2-6

The histogram for Cycle 2 is especially revealing as the MSD scores are significantly skewed to the right, with the majority of the sample scoring in top third of the score range. The suitability of the MSD for measuring cognitive ability was reviewed by Morongiello (1997) who found that "the MSD apparently 'tops out' (reaches ceiling) for children approaching three years, thereby not providing a sensitive test for these older children" [p.25]. The distribution of scores, along with Morongiello's findings, indicate that for the children in the present subsample, the MSD does little to differentiate among most of the children, and instead just serves as a screener for those experiencing significant delays.

For Cycles 3–6, the cognitive scores show a much wider distribution with the reported scores following a more normal distribution. The PPVT results are skewed slightly to the right and the CAT appears to cluster around certain scores. Though the distribution of scores is not perfectly normal, these tests show significantly more variation than the MSD scores from Cycle 2.

Using NLSCY Scores in this Thesis: Below, I discuss the analytical decisions that I have made so that the present analysis can make the most of the available cognitive scores in the NLSCY.

The first consideration is in regard to the relevance of the available cognitive measures. It is not enough to include a measure in my analysis simply because it exists in the data. To avoid the bias created by the significant ceiling on the MSD scores in Cycle 2, the final analysis does not include the MSD in the model of skill development. Due to the nature of the model, excluding the MSD requires omitting the entire stage from the full model and starting the model at Cycle 3.[129]

Once the relevant measures have been identified, the second consideration is choosing which form of the cognitive scores to include in my analysis. As mentioned at the beginning of this section, the cognitive scores are included in the analysis as z-scores. Specifically, my analysis uses what the NLSCY refers to as 'raw scores', for the PPVT; and 'classical scaled scores', for the CAT — but converts these scores to z-scores. This approach allows me to adjust for age separately within my model; therefore avoiding any bias from the norming sample used by the NLSCY to obtain their standardised scores.

Converting the scores to z-scores also allows me to maximise the size of the analysis sample by combining the measures from two tests taken in Cycle 4. Combining these scores is necessary because for Cycle 4 cognitive scores, my analysis sample is split across the age-boundary dividing two separate cognitive measures. During this NLSCY Cycle, the 6-year-olds were administered the PPVT, while the 7-year-olds were administered the CAT. If I were to use only one of these tests, the analysis sample would be greatly reduced. By converting the scores to z-scores, it is possible to combine both scores as a single rank-order measure for this age group. To control for this approach, all analysis includes an indicator variable for the type of assessment used to generate the Cycle 4 cognitive score.

[129] Appendix B.7 shows the results for an expanded model which includes Cycle 2. While some significance is lost as a result of expanding the model to include an additional 40 variables, the structure remains relatively unchanged.

Non-Cognitive Measures

As in the first empirical chapter, the term ‘non-cognitive’ is used in this empirical study as it is defined by Cunha and Heckman’s skill formation model. This means that non-cognitive ability refers to behavioural, social and emotional abilities. Given the more flexible definition of ‘non-cognitive’ when compared with some other measures, care must be taken to select the most relevant variables within the available data.

Unfortunately, it is difficult to compare behaviour between the two surveys directly as they do not share an identical set of measures. Unlike the MCS which used a pre-existing behavioural assessment in the form of the SDQ, the NLSCY data includes behavioural measures that are unique to the study. While they lack the broad evidence base that comes with the extensive use of the SDQ, the behavioural measures in the NLSCY provide significantly richer information as they draw on over 50 questions about behaviour in each cycle, compared to the 25-item SDQ.

The NLSCY behaviour questions are drawn from a variety of existing scales and surveys. Using these questions Statistics Canada has constructed composite indices which are known as the *NLSCY Behaviour Scale*. This scale reports a child’s behaviour using seven distinct factors, each corresponding to a known theoretical construct of behaviour. These constructs are: hyperactivity-inattention, emotional disorder-anxiety, physical aggression, separation anxiety, indirect aggression, property offences, and pro-social behaviour. Fortunately, many of the constructs used by the NLSCY correspond with those measured by the SDQ, so comparisons can be made between the two scales.[130]

Though these factor scores can be treated as stand-alone measures, a brief description of what they are designed to measure allows for better understanding of how they fit within my model. The design of the behaviour scale is outlined below, with a more detailed discussion provided in the User Guides for the NLSCY (Statistics Canada, 2004).

To create the behaviour scale, Statistics Canada convened *The NLSCY Expert Advisory Group* which included researchers from various fields relating to child development, as well as representatives from the federal and provincial departments involved in child development programmes in Canada. This group identified the types of behaviours they wanted to measure; and the relevant ages for measuring the selected behaviours. These decisions were motivated by the desire to create a survey instrument which would provide suitable measures of child development to allow for ongoing research, policy proposals; and which would meet the needs of relevant stakeholders. The original target behaviours were hyperactivity, emotional disorder, anxiety, physical aggression, inattention, prosocial behaviour, separation anxiety, opposition, indirect aggression and conduct disorder.

After choosing the target behaviours, Statistics Canada identified existing measures for them. This resulted in a list of questions from: the Ontario Child Health Study

[130] This overlap can be seen by comparing the list of questions from the SDQ presented in Table 4.7 with those from the NLSCY Behaviour Scale presented in Table 5.7.

(OCHS); the Montreal Longitudinal Study; Achenbach's Child Behavior Checklist (CBCL); a pro-social behaviour scale from Weir and Duveen (1981); and a measure of direct and indirect aggression from Lagerspetz, Björkqvist, and Peltonen (1988). These selected questions were included in the first cycle of the NLSCY.

Using the responses from Cycle 1 of the NLSCY, Statistics Canada used principal component analysis (PCA) to assess the psychometric properties of these behavioural measures. Although the questions correspond with an existing list of constructs, Statistics Canada wanted to confirm that the responses from the NLSCY corresponded with the factor structure of the theoretical constructs. This factor analysis found that the items loaded onto a slightly smaller group of behaviours than proposed by the advisory group, but the structures were generally in line with the proposed theoretical constructs.

To align with the previous chapter, which used the scale provided by the MCS data, I have chosen to accept the factor structure for the NLSCY behaviour scales. For more information about the factor structure, the reader is directed to the NLSCY user guides which provide results for the factor analysis, as well as a detailed discussion of the methods used. Though it might be possible to allow for a slight variation in the factor loadings within my subsample of choice, it is more desirable to maintain consistency with existing literature than to change the constructs to marginally improve model fit.

The final set of factors used by the NLSCY is presented in Table 5.7, along with the items used to construct each factor. The most notable variation is between Cycle 3 and Cycle 4, with the the measures of hyperactivity and emotional disorder having additional questions in Cycle 3. Though the factor structure remained relatively stable from Cycle 3 onwards, changes in the wording of some questions did occur between Cycle 3 and Cycle 4; however, these changes were minor and are listed in the footnotes of Table 5.7.

The behaviour questions are all included in the primary respondent portion of each cycle of the NLSCY. Though the questions are taken from different sources, the frequency of the behaviours described are all reported using the same scale. The PMK was asked how often each of the behaviours listed above described their children, and for each behaviour was given the choice of "never or not true", "sometimes or somewhat true" or "often or very true". The statements as included in the NLSCY questionnaire are provided in Table 5.7, though the survey presented the questions in a different order than that shown in the subscales listed here.

Separate scores were determined for each of the factors described in Table 5.7. Each behaviour is assigned a score which awards 0, 1 or 2 points for each of "never or not true", "sometimes or somewhat true" or "often or very true", respectively. Each of the behavioural subscales are scored as the sum of the scores on the relevant behaviour questions. Since the number of items differed between scales, the maximum value for each score reflects the number of items included in the scale. Higher scores correspond with

Table 5.7 Non-Cognitive Measures: NLSCY Behaviour Scales

Scale		How often would you say that your child...	C3	C4	C5	C6
Hyperactivity -Inattention	–	Can't sit still, is restless or hyperactive?[1]	X	X	X	X
	–	Is distractible, has trouble sticking to any activity?[2]	X	X	X	X
	–	Fidgets?	X			
	–	Can't concentrate, can't pay attention for long?	X	X	X	X
	–	Is impulsive, acts without thinking?	X	X	X	X
	–	Has difficulty awaiting turn in games or groups?	X	X	X	X
	–	Cannot settle to anything for more than a few moments?	X	X	X	X
	–	Is inattentive?	X	X	X	X
Pro-Social Behaviour	–	Will try to help someone who has been hurt?	X	X	X	X
	–	Shows sympathy to someone who has made a mistake?	X	X	X	X
	–	Volunteers to help clear up a mess someone else has made?	X	X	X	X
	–	If there is a quarrel or dispute, will try to stop it?	X	X	X	X
	–	Comforts a child (friend/sibling) who is crying or upset?	X	X	X	X
	–	Spontaneously helps pick up objects somebody has dropped?	X	X	X	X
	–	Will invite others to join in a game?	X	X	X	X
	–	Helps other children (friends/siblings) who are feeling sick?	X	X	X	X
	–	Helps those who do not do as well as he does?	X	X	X	X
	–	Offers to help other children (friends/siblings) who are having difficulty with a task?	X	X	X	X
Emotional Disorder-Anxiety	–	Seems to be unhappy, sad or depressed?[3]	X	X	X	X
	–	Is not as happy as other children?	X	X	X	X
	–	Is too fearful or anxious?	X	X	X	X
	–	Is worried?	X	X	X	X
	–	Cries a lot?	X	X	X	X
	–	Appears miserable, unhappy, tearful, or distressed?	X			
	–	Is nervous, high strung or tense?	X	X	X	X
	–	Has trouble enjoying him/her self?	X	X	X	X
Physical Aggression	–	Gets into many fights?	X	X	X	X
	–	Physically attacks people?	X	X	X	X
	–	Threatens people?	X	X	X	X
	–	Is cruel, bullies or is mean to others?[4]	X	X	X	X
	–	Kicks,bites or hits other children?	X	X	X	X
	–	When somebody accidentally hurts him/her, he/she reacts with anger and fighting?	X	X	X	X
Indirect Aggression	When mad at someone, ...					
	–	... tries to get others to dislike that person?	X	X	X	X
	–	... becomes friends with another as revenge?	X	X	X	X
	–	... says bad things behind the other's back?	X	X	X	X
	–	... says to others: let's not be with him/her?	X	X	X	X
	–	... tells that person's secrets to a third person?	X	X	X	X
Property Offenses	–	Destroys his/her own things?	X		X	X
	–	Steals at home?	X		X	X
	–	Destroys things belonging to his/her family, or other children?	X		X	X
	–	Tells lies or cheats?	X		X	X
	–	Vandalizes?	X		X	X
	–	Steals outside his/her home?	X		X	X

Notes:
[1] The word hyperactive was removed for C4, C5 and C6.
[2] Distractible was replaced with 'easily distracted' in C4, C5 and C6.
[3] The word depressed was removed for C4, C5 and C6.
[4] 'Is cruel' was removed for C4, C5 and C6.

higher frequencies of the behaviours indicative of hyperactivity-inattention, emotional disorder/anxiety, physical aggression, separation anxiety, indirect aggression and property offences, while higher pro-social scores correspond with a higher incidence of pro-social behaviours.

Included in the NLSCY data are the responses for each individual reported behaviour. In line with the previous chapter, I have chosen to use the derived subscale scores within my research. Using the scores for the specific scales meets three objectives. First, using the accepted scales allows me to compare the results of my study with other research using the NLSCY data. As the present study aims to investigate the impact of parental investment, it is best to use existing measures and allow the focus to remain on the results which can be discussed within the context of other research. Secondly, the confidentiality requirements of the NLSCY require that any variables used in the analysis do not identify a group of fewer than ten children. As the number of variables used increases, the probability of an interaction term of two variables violating this restriction grows proportionately. Thirdly, and perhaps most importantly, using the individual behaviours adds over 200 variables to my model, substantially reducing the ability to obtain statistically significant results.[131]

The summary statistics for the SDQ are presented in Table 5.8. Hyperactivity scores decrease over time, with the standard deviation remaining relatively constant. This matches the pattern seen in the hyperactivity scores in the MCS. Physical aggression appears to decrease over time, as do property offences. The questions from the hyperactivity and physical aggression scales map closely onto those used in the SDQ scale for conduct problems: the behaviours these questions describe show a similar decrease over time in both the NLSCY and the MCS. There is a slight increase in pro-social behaviour as the sample ages, which is larger than the change seen in pro-social behaviour in the MCS. Finally, symptoms of emotional disorders in the NLSCY data show no clear trend.

Table 5.8 Measures of Non-Cognitive Ability: Weighted Descriptive Statistics

		Cycle 3 *Age 4/5*		Cycle 4 *Age 6/7*		Cycle 5 *Age 8/9*		Cycle 6 *Age 10/11*	
	range[1]	mean	SD	mean	SD	mean	SD	mean	SD
Hyperactivity-Inattention	0–14	4.204	2.898	4.075	2.975	3.994	2.992	3.442	3.045
Pro Social behaviour	0–20	7.794	3.908	5.920	3.521	5.671	3.685	5.298	3.631
Emotional Disorder-Anxiety	0–14	1.830	1.842	2.228	2.103	2.603	2.204	2.396	2.247
Physical Aggression	0–12	1.669	1.879	1.451	1.896	1.356	1.886	1.018	1.634
Indirect Aggression	0–10	0.622	1.200	0.953	1.477	0.990	1.472	0.911	1.546
Property Offences	0–12	1.165	1.471			0.972	1.290	0.632	1.029

Notes:

[1] Each behaviour scale was based on a different number of questions. Thus, the total score for each scale varied.
Note: All descriptive statistics are calculated using provided survey weights and bootstrap weights.

[131] As discussed above, detailed information about this factor structure is provided by Statistics Canada in the relevant user guides.

Though the present analysis uses the behavioural scores as provided, there was one small adjustment made to the scores for Cycle 3 to allow for consistency across all four cycles. The scores for ‘hyperactivity-inattention’ and ‘emotional disorder-anxiety’ were re-calculated using the same scale that used for Cycles 4–6. More specifically, in Cycle 4 the question “How often would you say that your child is inattentive?" was excluded from the hyperactivity-inattention score and “How often would you say that your child appears miserable, unhappy, tearful or distressed?" was omitted from the emotional disorder-anxiety score. Analysis was conducted using both versions of these scores and no significant differences were found. [132]

Histograms of the scores are shown in Figure 5.4, Figure 5.5, Figure 5.6 and Figure 5.7. These histograms show that none of the behaviour variables exhibited normal distributions, with significant left clustering for all variables. As with the MCS, this distribution is to be expected as the scales are derived from measurements of psychopathology in children.

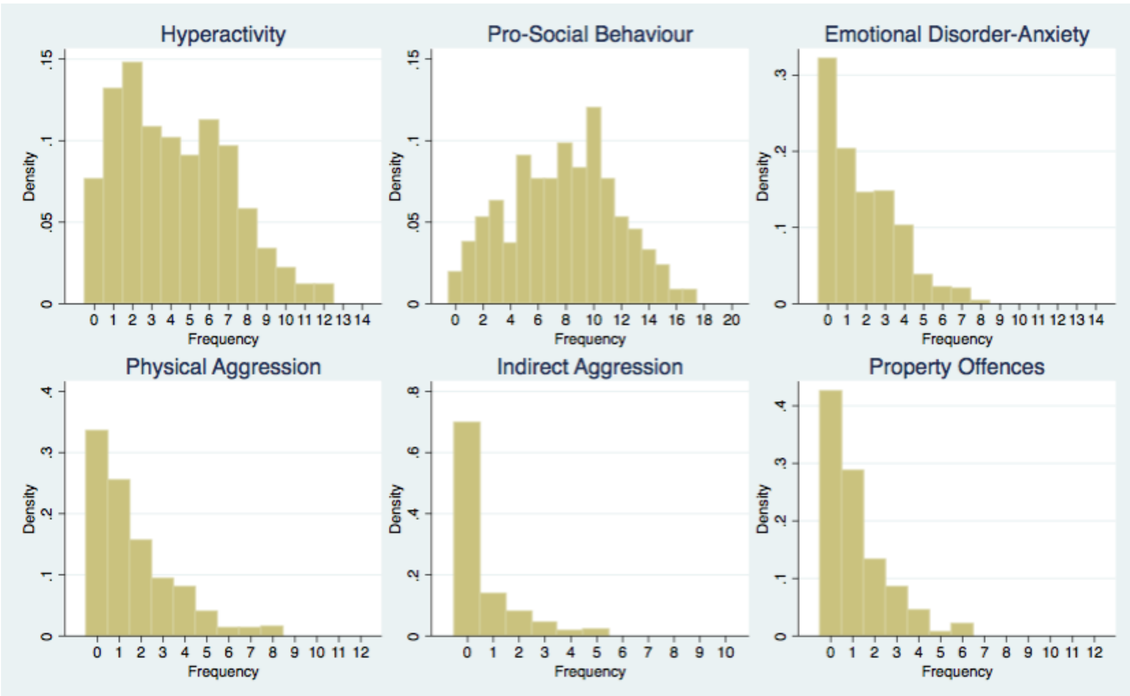


Fig. 5.4 Distribution of Non-Cognitive Scores: NLSCY Cycle 3 (Ages 4/5)

[132]To comply with Statistics Canada regulations, the results from this additional analysis are not presented in this thesis.

Fig. 5.5 Distribution of Non-Cognitive Scores: NLSCY Cycle 4 (Ages 6/8)

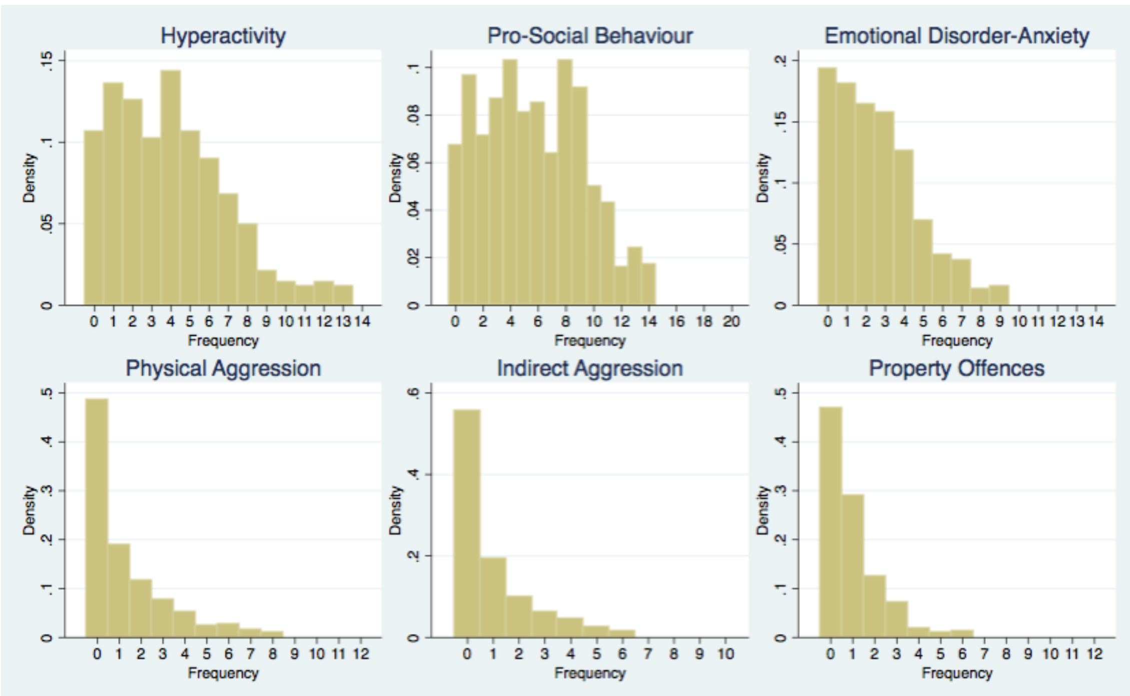


Fig. 5.6 Distribution of Non-Cognitive Scores: NLSCY Cycle 5 (Ages 8/9)

Fig. 5.7 Distribution of Non-Cognitive Scores: NLSCY Cycle 6 (Ages 10/11)

Investment Measures

As with non-cognitive ability, the choice of parental investment measures is strongly influenced by the questions included within the NLSCY survey design. In line with the UK empirical study, the term parental investment is used to describe direct interaction between parents and their children. I aimed to identify behaviours that required time to be spent directly with the child, and activities that had no significant financial cost. The most suitable measure for this was included in the NLSCY as the *NLSCY Parenting Scale* which consisted of 25 questions measuring parent-child interactions. According to Statistics Canada (2010), this portion of the survey was designed "to measure the positive interaction, hostility/ineffectiveness, and consistency of parenting the child" as well as "parental practices that may or may not provoke aversion" in the child [p.79].

The questions used in the parenting scales were provided by Dr. M. Boyle, a child development researcher, who was involved with survey design for Statistics Canada. Dr. Boyle provided 25 questions about parenting, with the first 18 asking about the frequency with which the PMK engages in various behaviours with their child and the remaining seven asking about how the parent responds when the child breaks rules. The first 18 questions were an adaptation of the Parent Practices Scale introduced by Strayhorn and Weidman (1988), while the remaining seven were based on Dr. Boyle's own research.[133]

To provide some international comparability and ability to compare outside the NLSCY sample, I have chosen to focus on the first 18 of these questions. In addition to their lack of external validity, the final seven parenting questions focus on style of punishment and are distinct from the parenting behaviours described by the other behavioural questions. The 18 questions that are used in the analysis are listed in Table 5.9 and are repeated across all cycles of the NLSCY.[134]

Using the scores collected for the Parenting Scales questions, Statistics Canada conducted its own factor analysis of the 18 parenting measures. This factor analysis found that 17 of the questions loaded onto three separate parenting factors: positive interaction; hostile/ineffective interaction; and consistent parenting. For each of these three factors, a factor score was created by summing the individual scores of the behaviours that loaded onto that factor. Though it would have been possible to use these equal-weighted factor scores, I chose to leave the variables in their raw form and conduct my own factor analysis. This factor analysis is discussed in greater detail in the results section. Furthermore, the parenting measures are left as scores from 1–5 which allows me to ensure that their

[133]The first 18 questions are an excerpt of the 27 questions included in the Parent Questionnaire that is used widely by the Fast Track program in the US.

[134]The PMK's responses are coded on one of two scales. The first scale is set as: 1 – "never"; 2 – "about once a week or less"; 3 – "a few times a week"; 4 – "1–2 times a day"; 5 – "many times each day". The second frequency scale is defined as: 1 – "never"; 2 – "less than half the time"; 3 – "about half the time"; 4 – "more than half the time"; 5 – "all the time". These response scales result in a score of 1–5 for each of the questions listed.

eventual weighting be determined within the structural equation model used for my analysis.

Table 5.9 Parental Investment Measures: NLSCY Parenting Scale Questions

	Question
1.	How often do you praise [your child], by saying something like "Good for you!" or "What a nice thing you did!" or "That's good going!"?
2.	How often do you and [your child] talk or play with each other, focusing attention on each other for five minutes or more, just for fun?
3.	How often do you and [your child] laugh together?
4.	How often do you get annoyed with [your child] for saying or doing something he/she is not supposed to?
5.	How often do you tell [your child] that he/she is bad or not as good as others?
6.	How often do you do something special with [your child] that he/she enjoys?
7.	How often do you play sports, hobbies or games with [your child]?
8.	Of all the times that you talk to [your child] about his/her behaviour, what proportion is praise?
9.	Of all the times that you talk to [your child] about his/her behaviour, what proportion is disapproval?
10.	When you give [your child] a command or order to do something, what proportion of the time do you make sure that he/she does it?
11.	If you tell [your child] they will get punished if he/she doesn't stop doing something, and he/she keeps doing it, how often will you punish him/her?
12.	How often does he/she get away with things that you feel should have been punished?
13.	How often do you get angry when you punish [your child]?
14.	How often do you think that the kind of punishment you give [your child] depends on your mood?
15.	How often do you feel you are having problems managing [your child] in general?
16.	How often is [your child] able to get out of a punishment when he/she really sets his/her mind to it?
17.	How often when you discipline [your child], does he/she ignore the punishment?
18.	How often do you have to discipline [your child] repeatedly for the same thing?

Notes:
[1] In the survey questionnaire [your child] was replaced with the child's name.
[2] Questions 1-7 are reported using: 1 – "never"; 2 – "about once a week or less"; 3 – "a few times a week"; 4 – "1–2 times a day"; 5 – "many times each day".
[3] Questions 8-18 are reported using: 1 – "never"; 2 – "less than half the time"; 3 – "about half the time"; 4 – "more than half the time"; 5 – "all the time" .

Although the NLSCY Parenting Scales are an adaptation of an existing scale and were not independently validated, they have been used in multiple studies since the first cycle of the NLSCY — including in studies that find relationships between the parenting scales and a variety of outcomes in childhood and adolescence. Landy and Tam (1996) find that positive parenting predicts higher scores on measures of social development. Building on their initial findings, Landy and Tam (1998) examine the relationship between parenting scales and the child behaviour measures of emotional disorder, hyperactivity, aggressive behaviour, and conduct disorder. These findings will be discussed in greater detail in relationship to the results from my own analysis, but the use of the parenting scales by Landy and Tam, along with other work by Chao and Willms (2000), provide support for the use of the NLSCY measures of parenting.

Summary statistics are presented in Table 5.10. The first six variables appear to decrease in frequency as the child gets older. The remaining variables follow a mixed pattern of change over time. Some, such as 'ability to avoid punishment', remain relatively constant while others, such as 'ignoring punishment', rise over time. This distinctive pattern of change indicates a possible flaw with using the NLSCY constructed factors because an increase in one variable would be cancelled out by an increase in another. By estimating the factor loadings independently for each cycle, a more detailed understanding of development will be possible.

Table 5.10 Parental Investment Measures: Weighted Descriptive Statistics

	Cycle 3 *Age 4/5*		Cycle 4 *Age 6/7*		Cycle 5 *Age 8/9*	
	mean	SD	mean	SD	mean	SD
Praise child[1]	4.483	0.678	4.293	0.742	4.069	0.771
Five minutes focused attention[1]	4.230	0.718	4.004	0.775	3.856	0.803
Laugh together with child[1]	4.605	0.605	4.426	0.682	4.298	0.725
Get annoyed with child[1]	3.329	0.958	3.126	0.936	3.093	0.948
Tell child they are bad[1]	1.251	0.645	1.125	0.409	1.178	0.489
Do something special with child[1]	3.101	0.873	2.718	0.732	2.611	0.728
Play sports with child[1]	3.066	0.826	2.760	0.734	2.663	0.709
Proportion of talk: praise[2]	2.135	0.631	2.236	0.748	2.268	0.736
Proportion of talk: disapproval[2]	2.185	0.608	2.219	0.664	2.258	0.642
Make sure child follows commands[2]	4.191	0.857	4.312	0.781	4.376	0.757
Punish child if breaks rules[2]	3.881	1.149	3.989	1.105	4.036	1.112
Child gets away with breaking rules[2]	3.934	0.828	3.946	0.851	4.006	0.780
Angry while punishing child[2]	2.239	0.918	2.151	0.892	2.128	0.871
Punishment depends on mood[2]	2.167	1.077	2.124	1.026	2.030	0.984
Problems managing child[2]	1.703	0.797	1.622	0.778	1.624	0.804
Child able to avoid punishment[2]	3.829	1.121	3.820	1.123	3.840	1.098
Child ignores punishment[2]	4.143	0.977	4.313	0.895	4.377	0.854
Need for repeated discipline[2]	2.294	0.938	2.124	0.852	2.071	0.861

Notes:
– All descriptive statistics are calculated using provided survey weights and bootstrap weights.
– All parenting measures are reported on five-point scale:
[1] Coded using: 1 – "never"; 2 – "about once a week or less"; 3 – "a few times a week"; 4 – "1–2 times a day"; 5 – "many times each day".
[2] Coded using: 1 – "never"; 2 – "less than half the time"; 3 – "about half the time"; 4 – "more than half the time"; 5 – "all the time" .

Figure 5.8, Figure 5.9 and Figure 5.10 show the categorical distributions of the parental investment measures. The histograms provide slightly more insight into how these behaviours are distributed within the sample.

At ages 4–5, as shown in Figure 5.8, many of the variables have few responses for one of the two extreme responses (i.e. ‘all the time’ and ‘never’). That is, very few parents report a total absence of praise, laughing with their child, doing special things with the child, playing sports with the child or spending five minutes focusing on the child. Similarly, there are few parents who report that they have difficulty managing their child, speak with disapproval or require the use of repeated discipline: ‘all the time’.

At ages 6–7, as shown in Figure 5.9, the parental investment measures are largely similar with slight shifts in the relative distribution. Compared to ages 4–5, many of the histograms at ages 7–9 tend to shift their distribution to the left. The exceptions are the right shift in the reported frequency of laughing with the child, and the frequency of experiencing anger when punishing the child.

Again, at ages 8–9, as shown in Figure 5.10, the distributions appear to be slightly less skewed towards the right with a broader range of responses for some variables. This indicates that as the child ages, parents vary more in the frequency of specific parenting behaviours.

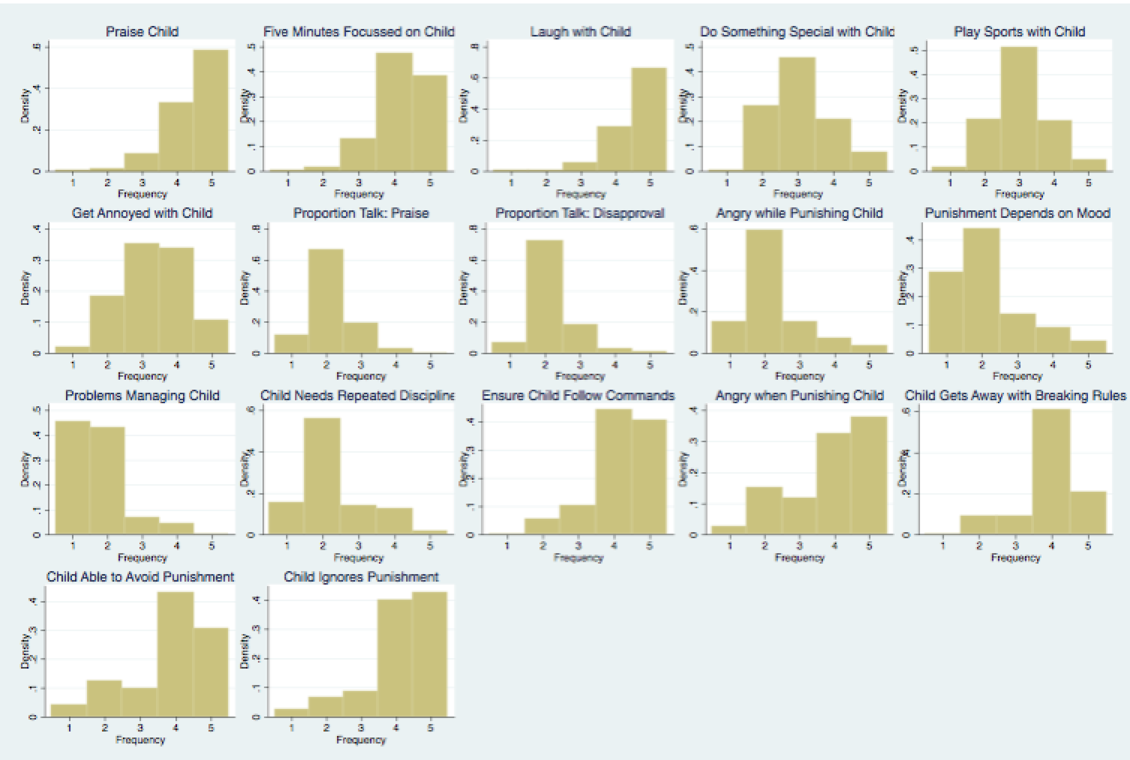


Fig. 5.8 Distribution of Parental Investment Measures: NLSCY Cycle 3 (Ages 4/5)

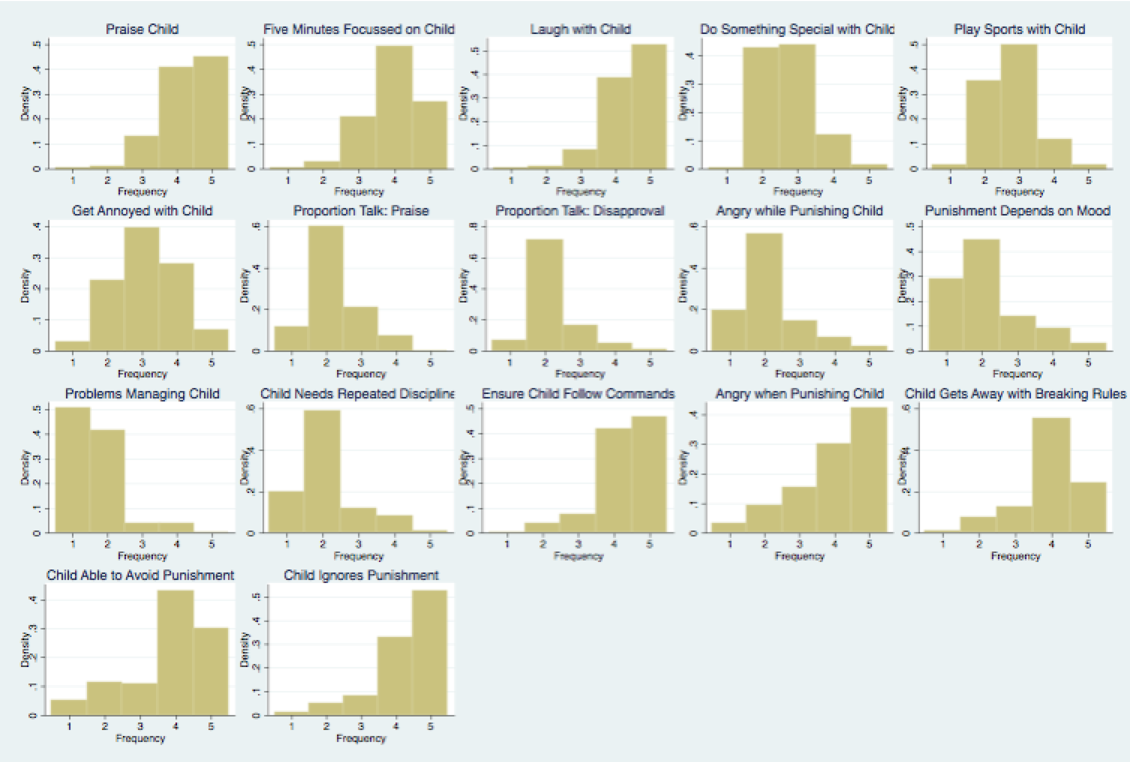


Fig. 5.9 Distribution of Parental Investment Measures: NLSCY Cycle 4 (Ages 6/7)

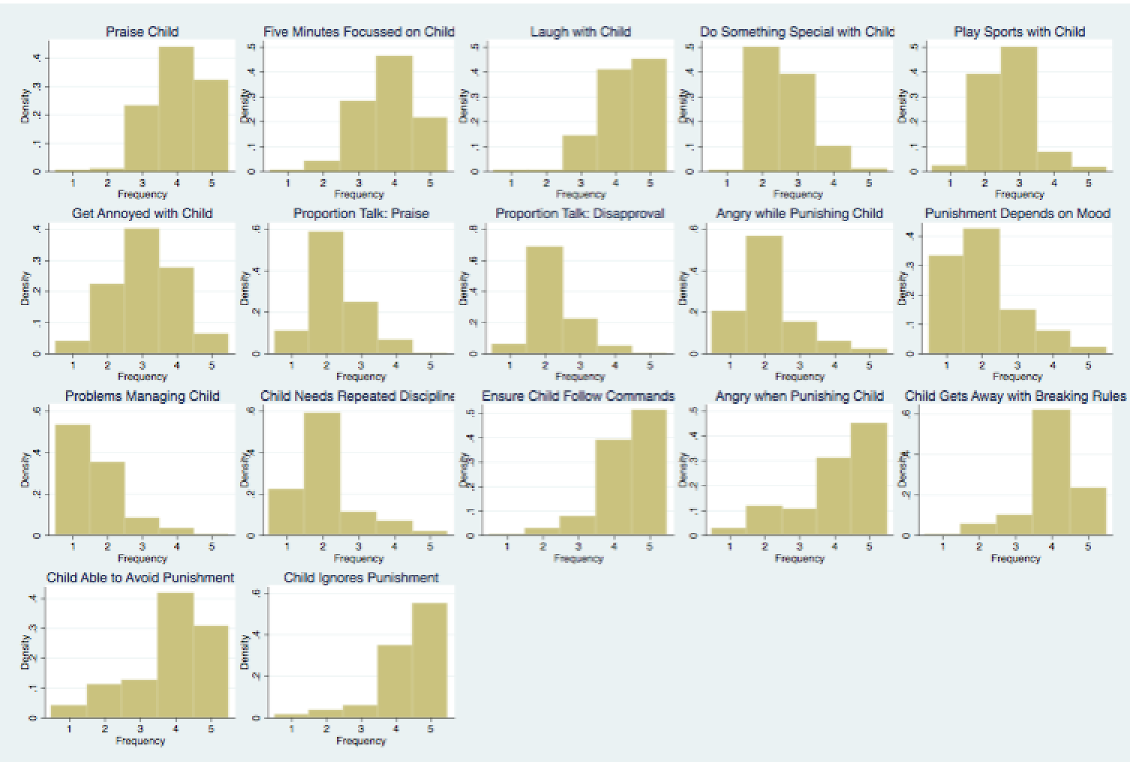


Fig. 5.10 Distribution of Parental Investment Measures: NLSCY Cycle 5 (Ages 8/9)

5.3 Empirical Model

Using the NLSCY data described above, it is possible to estimate the dynamic model presented in Chapter 3. As with Chapter 4, this chapter does not restate the empirical methodology, but instead reviews the key details in the context of the Canadian data. Although this section focuses on the Canada-specific application of the model, some parts of the text are repeated verbatim from the previous chapter. This repetition is necessary for context within the present chapter and to save the reader from having to return to Chapter 4 to search for the relevant equations for this application.

Structural Model

As before, a child's current ability θ_t consists of cognitive and non-cognitive skills: $\theta_t = \left(\theta_t^C, \theta_t^N\right)'$. At any point in time $(t+1)$, this set of skills is a function of: a child's skill in the last period — θ_t; parental behaviour (i.e. parental investment) — I_t; and observable exogenous measures of socio-economic status — X_t. A recursive, linear function can be used to express the evolution of skill over time as

$$\theta_{t+1} = \Gamma_t\theta_t + B_t I_t + \Lambda_t X_t + \eta_t, \tag{5.1}$$

for $t \in 1, \ldots, T$, where θ_t is a latent vector of skills; I_t is an $(s \times 1)$ latent vector of parental inputs[135]; X_t is an observed matrix of exogenous variables; and η_t is the error term. In the Canadian context, X_t contains the categorical variable measuring household income quintile, while the definitions for θ_t and I_t are provided later in this section.

Each of the s types of parental investment are influenced by the matrix of observable variables X_t^I, and so vector of parental investments can be expressed by the function

$$I_t = \phi_t^I X_t^I + \varsigma_t, \tag{5.2}$$

where ϕ_t^I is a matrix of estimated parameters; and ς_t is the error term. As with the first empirical application, in the Canadian model, the matrix X_t^I captures single parent household status and number of siblings.

A child's skill development must have an initial starting point, and it is assumed that a child begins life with an initial endowment of skill θ_0, modelled as

$$\theta_0 = \psi_0 X_0^\theta + \xi_0, \tag{5.3}$$

where ψ_0 is a matrix of estimated parameters; and X_0^θ contains both the period-specific measures, included in X_t, and time-invariant demographic characteristics used to capture family background and early child health. As with the MCS, the relevant measures of X_0^θ from the NLSCY are birth weight, mother's age at birth, maternal education, and household income quintile. The specifics of these measures are presented in Table 5.3.

[135]By defining I_t as a vector (ie. $s \neq 1$), this structural model allows the possibility of multiple types of parental investment.

Though the full model is again estimated simultaneously, it can be reduced to two separate linear laws of motion for non-cognitive and cognitive skills:

$$\theta_{t+1}^{C} = \gamma_{1,t}^{C}\theta_t^C + \gamma_{2,t}^{C}\theta_t^N + \beta_{1t}^{C} I_t^1 + \cdots + \beta_{st}^{C} I_t^s + \Lambda_t^C X_t + \eta_t^C \tag{5.4}$$

and

$$\theta_{t+1}^{N} = \gamma_{1,t}^{N}\theta_t^C + \gamma_{2,t}^{N}\theta_t^N + \beta_{1t}^{N} I_t^1 + \cdots + \beta_{st}^{N} I_t^s + \Lambda_t^N X_t + \eta_t^N \tag{5.5}$$

respectively. Though both laws of motion contain the same variables, the estimated parameters are specific to each type of skill.

Measurement Model

Unlike in Chapter 4 where the MCS data allows me to use measurement models to estimate cognitive skills, non-cognitive skills and parental investment, the NLSCY data only contains sufficient measures to create latent measurement models for non-cognitive ability and parental investment. The NLSCY-specific indicators and measurement models for each of these latent factors are presented below, along with an explanation of the alternative approach to measuring cognitive ability.

Cognitive Skills

Since the NLSCY contains only one measure of cognitive skill per cycle, I am unable to use a measurement model and must directly include cognitive ability in the larger structural model. Thus, the chosen measures of cognitive ability correspond with θ_t^C in Equation 5.4 and Equation 5.5.

As the data does not allow for the use of a measurement model, inherent measurement error from the cognitive measures is assumed. This measurement error can be partially corrected for by including control variables in Equation 5.4 for factors known to cause measurement error for cognitive ability. In the case of the NLSCY, these control variables are gender, age at time of test and language spoken in the home. Though this is not ideal, the cognitive measures are well designed and still allow for conceptual proof of the model within a Canadian context. Fortunately, the cognitive measures contained in the NLSCY are widely used, so information is available on how representative they are of true underlying cognitive skill.

Non-Cognitive Skills

The NLSCY behavioural scales contains multiple observable indicators of non-cognitive skills, with details presented in Table 5.8. The number of available behaviour scales in a cycle varies and is given by m_t^N, with $m_t^N = 5$ or $m_t^N = 6$. Therefore, the set of indicators for each cycle (N) is expressed as $Y_{j,t}^N$, $j \in \{1, \ldots, m_t^N\}$.

Assessing the Factor Structure: Unlike the SDQ measures used in the first empirical chapter, the behavioural measures in the NLSCY have not been previously validated as identifying one latent construct. This means that while the previous chapter confirmed the reliability of the scale, this chapter calculates Cronbach's alpha to *assess* the reliability of combining the items in a single scale. Additionally, by calculating separate values of Cronbach's alpha for the test scale with each variable removed, it is possible to identify poorly fitting items. If the removal of any behaviour subscale significantly raises the value of α then it may not identify the underlying construct. If this is the case, the decision must be made whether or not to include this scale in predicting the underlying factor.

Applying the Measurement Model: Using the subscales that have been identified as measuring non-cognitive ability, I am able to estimate a score for the latent factor. Though it would be easiest to take the sum of the scores on these behavioural measures, it is unlikely that each measure captures the same amount of information about the true underlying level of skill. As with the SDQ subscales in the previous chapter, I use a measurement model to estimate a latent score for non-cognitive ability.

In this measurement model, any given behaviour subscale $(Y^N_{j,t})$ is given by

$$(Y^N_{q,t})^* = \mu^N_{q,t} + \alpha^N_{q,t}\theta^N_t + \Phi^N_{q,t}Z^N_t + \varepsilon^N_{q,t}, \tag{5.6}$$

such that $Y^N_{q,t} = r$ if $\rho^q_{r-1} \leq (Y^N_{q,t})^* \leq \rho^q_r$ where $\rho^q_{r-1} = -\infty$ and $\rho^q_{R^q_t} = \infty$. There is one equation for each of the m^N_t behaviour scores for the given cycle. To model some of the measurement error caused by observable variables, each equation includes the matrix of covariates Z^N_t that are known to influence the measured indicator but are independent of the underlying latent factor. I have previously outlined how parental reports of behaviour are known to differ systematically based on the child's gender, ethnicity and age. As with the measurement model in Chapter 4, these three factors are included as Z^N_t in the measurement models for this chapter. Defining Equation 5.6 for each of the m^N_t behaviour scales in a given NLSCY cycle allows me to estimate the child's latent non-cognitive ability for that cycle.

Parental Investment

In both Chapter 4 and this chapter, the model assumes the existence of $S \geq 1$ different types of unobservable parental investment, but the while cognitive and non-cognitive ability are consistently defined across the two empirical applications, the types of parental investment are strongly driven by the available data. Since one goal of this thesis is to demonstrate the adaptability of the skill formation model, it is important to show that the choice of parenting indicators need not match from one empirical application to the next. The methodology that I use to define and measure parental investment in the Canadian context is outlined below.

Defining the Factor Structure: Before applying the measurement model, the separate types of parental input must be identified. In the empirical application using the MCS data, I identified a set of similar behaviours in the data, and then used these to define the parenting constructs. For this chapter, I take a slightly different approach and use a set of parenting constructs already identified by Statistics Canada. Specifically, the NLSCY includes an assessment of parenting behaviours which measure pre-defined categories of parenting practices. The 18 questions that were included in the NLSCY were each designed to correspond to one of three measures of parenting defined by Strayhorn and Weidman (1988): warmth and involvement; consistency; and punitive disciplinary tactics. The first two of these categories share significant overlap with the literature on parenting style discussed in Section 2.4 — specifically the parenting style framework identified by Maccoby and Martin (1983) and Darling and Steinberg (1993).

Based on the available data, this empirical application models investment using measures of parenting style, and applying the methodological approach used in the previous chapter. Although it would be possible to simply separate the questions into the three types of parenting behaviours defined by Strayhorn and Weidman (1988), the NLSCY documentation notes that the parenting measures do not follow the expected factor structure. Statistics Canada (2004) provides results from their own factor analysis of Cycle 1 data for children ages 4–11 and proposes NLSCY-specific parenting categories.

While I do make note of the parenting categories proposed by the NLSCY, I use factor analysis to identify the structure within the analysis sample. Applying EFA to the 18 variables, I estimate the eigenvalue, RMSEA, SRMR, CFI and TLI and use the cutoff criteria described in Section 3.3.2 to confirm the number of latent investment factors (i.e. parenting styles) to be included in the measurement model.[136]

Applying the Measurement Model: After using the 18 parenting questions to identify the $s \in \{S \geq 1\}$ latent investment factors I_t^s, I create a measurement model for each latent factor. Each factor corresponds with a set of m_t^s parenting practices, with each practice expressed by $Y_{q,t}^s, q \in \{1, \ldots, m_t^s\}$. In the NLSCY, parenting practices are reported categorically as r =*1,2,3,4 or 5* categories, resulting in the measurement model:

$$(Y_{q,t}^s)^* = \mu_{q,t}^s + \alpha_{q,t}^s I_t^s + \Phi_{q,t}^s Z_t^s + \varepsilon_{q,t}^s, \tag{5.7}$$

such that $Y_{q,t}^s = r$ if $\rho_{r-1}^q \leq (Y_{q,t}^s)^* \leq \rho_r^q$ where $\rho_{r-1}^q = -\infty$ and $\rho_{R_t^q}^q = \infty$. There are m_t^s equations, one for each indicator. Each measurement equation includes the matrix of covariates Z_t^s that are known to influence the measurement of the given behaviour but are independent of the underlying latent factor. In the case of the NLSCY, there are no known demographic measures which correspond with systematic measurement error of the self-reported parenting practices, so Z_t^s is not included in the model.

[136] Earlier in this thesis, Section 3.3.1 provided a detailed explanation of EFA and the use of fit statistics. For the sake of brevity, these are not reproduced in this chapter.

Analysis Procedures

As with the MCS, the NLSCY is also provided as a set of separate data files. Using the unique person identifier[137], I was able to merge the appropriate files. The data was cleaned, and descriptive statistics were calculated in Stata 14.0. EFA and the full model were estimated in Mplus 8.0 (Muthén & Muthén, 2017). All reported values were calculated using the Statistics Canada provided survey weights to adjust for the original sampling strategy and attrition.[138] For reference, the MPlus input file for the full dynamic model is provided in Appendix B.6. Additionally, the provided bootstrap weights were used to calculate appropriate standard errors.[139] As these sample weights are based on the Canadian population in 1993 and do not account for immigration or emigration, they may not be representative of the present-day Canadian population.

[137]The NLSCY identifies each child with a unique person identifier, this variable is called "*PERSRUK*". This variable is consistent across cycles.

[138]The NLSCY provides three different sets of sample weights for each cycle of the NLSCY. This analysis uses the sample weights contained in the variable '*HWTCWd1L*', which are the longitudinal funnel weights from Cycle 8. For more information see the NLSCY User Guide.

[139]More information about the use of survey weights and bootstrapping is provided in Appendix B.5.

5.4 Results

Measurement Model: Cognitive Ability

As there is only one cognitive measure available for the relevant age group for each cycle of the NLSCY, there is no direct measurement model for cognitive skills. Instead, the cognitive measures are included directly as θ_t^C in Equation 5.4 and Equation 5.5. The justification for making this adjustment to the model has been discussed previously in the methodology section of this chapter. The full model does however include controls for the relevant covariates. Though this is not strictly a measurement model, the results show the extent to which the scores on cognitive ability are related to the control variables.

Table 5.11 Covariate Parameter Estimates: Cognitive Ability

	Cycle 3 *Age 4/5*	Cycle 4 *Age 6/7*	Cycle 5 *Age 8/9*	Cycle 6 *Age 10/11*
Female	-0.121***	-0.176***	-0.075*	0.000
	(0.038)	(0.040)	(0.041)	(0.047)
White	0.201	0.159	-0.438***	0.099
	(0.137)	(0.112)	(0.125)	(0.109)
No English/French Spoken In Home	-0.684***	-0.026	0.247*	-0.069
	(0.122)	(0.118)	(0.126)	(0.099)
Child Age (in months)	0.063*	0.022*	0.021	0.018
	(0.037)	(0.011)	(0.025)	(0.019)

Notes:
- All models are estimated using provided survey weights and bootstrap weights.
- *p<0.1 , **p<0.05, ***p<0.01
- Standard errors in parentheses.

To maintain consistency with the first empirical chapter, the coefficients on these parameters are presented with the measurement model and are shown in Table 5.11. As seen in the MCS data, these coefficients show that there are significant differences in cognitive scores for children of differing race, gender or native language.

Unsurprisingly, the child's age in months correlates with higher scores for Cycles 3 and 4. This is in line with the information provided with the NLSCY that suggests using age standardisation within any models using the unscaled cognitive scores. The effect of age appears to fall over time, indicating that earlier cognitive ability may change more from one month to the next. Age does not appear to correspond with test scores for Cycles 5 and Cycle 6. This might indicate that these tests are less sensitive to age at the time of testing.

For other covariates the impact varies between cycles. At ages 4–5 (Cycle 3), children who speak neither English nor French at home tend to score significantly lower than those who speak either official language, with non-native speakers having scores

0.684 standard deviations below the mean. The Cycle 3 PPVT is the only cognitive test that takes place before the child enters formal schooling. As the tests are conducted only in English or French, it might indicate that these children do not have sufficient language skills for the test to adequately measure their performance. For Cycles 4 and 6 the impact of language spoken in the home is non-significant, perhaps indicating that exposure to an official language in school allows for these tests to suitably measure these children's ability. Interestingly, at ages 8–9, children who are non-native speakers score 0.247 standard deviations higher than native speakers.

This pattern of achievement is consistent with UK test score data, where English as an additional language (EAL) children initially under-perform relative to their native-English-speaking peers, but outperform them as they grow older (see Demie, 2018; Strand, Malmberg, & Hall, 2015). In Chapter 4, this pattern was not as pronounced as in the Canadian data, though it was demonstrated by the lack of significance on the MCS coefficients for language status on the cognitive measures at age 11.

For the NLSCY data, the impact of race on test scores, after controlling for all other factors, appears to be largely insignificant — with the only observed effect being during Cycle 5, when the children are 8-9 years old. For this cognitive test, white children score 0.438 standard deviations below the mean. This result is striking because this difference disappears during Cycle 6, even though both Cycles 5 and 6 are based on the same CAT mathematics test.

Finally, in contrast to the UK sample in the previous chapter, where male respondents had lower cognitive ability, in the Canadian data female respondents have lower cognitive scores than their male peers. In a comparative analysis of Canada, the UK, the US and Australia, Bradbury, Waldfogel, Washbrook, and Corak (2015b, p.89) find a similar pattern of gender differences in the UK and Canadian data. Their findings confirm the pattern seen in my two empirical projects; whereby the impact of gender differs between the two countries but in both countries the magnitude of this effect reduces over time.

Measurement Model: Non-Cognitive Ability

Confirming the Factor Structure

The NLSCY behavioural scales have not been validated as a single construct, so the present analysis assesses reliability of a single non-cognitive scale. To do so, Cronbach's alpha was calculated separately for each cycle and the results are presented below in Table 5.12.

Table 5.12 Measures of Non-Cognitive Ability: Scale Reliability

	Raw Cronbach's Alpha	Item that lowers Cronbach's Alpha most if excluded	Cronbach's Alpha if the Item is Excluded
Cycle 3 - *Age 4/5*	0.643	Pro-Social Behaviour	0.698
Cycle 4 - *Age 6/7*	0.609	Pro-Social Behaviour	0.644
Cycle 5 - *Age 8/9*	0.704	Pro-Social Behaviour	0.727
Cycle 6 - *Age 10/11*	0.731	Pro-Social Behaviour	0.755

Notes:

- Cronbach's alpha was calculated for the scale omitting each of the included variables.
- Full results are omitted to comply with Statistics Canada regulations.

In Table 5.12, one can see that removing pro-social behaviour raises the α in all cycles. This indicates that including the score for pro-social behaviour decreases the reliability of the non-cognitive scale. To confirm this finding, EFA was conducted for the non-cognitive measures. Though the full findings are not reported in this thesis in order to comply with Statistics Canada regulations, they point towards a two-factor solution with pro-social behaviour loading onto a single factor while the remaining behaviour scales load onto another factor. The results from these two assessments of fit led me to drop the measure of pro-social behaviour from the non-cognitive factor scale. This might reflect the fact that the absence of pro-social behaviours, which are positive, is conceptually different from the presence of negative behaviours.[140]

Results from the Measurement Model

The parameter estimates for the non-cognitive measurement model are shown in Table 5.13. Panel A provides the estimates for $\alpha^N_{j,t}$ in Equation 5.6. The first factor loading is set as $\alpha^N_{1,t} = -1$ to simplify the interpretation of the structural model. This results in all the other factor loadings taking negative values. Higher levels of non-cognitive development correspond with lower parental reports of hyperactivity, emotional symptoms, physical aggression, indirect aggression and property offences. The magnitude of the factor loadings changes from one period to the next, indicating that each of the measures contains a different amount of information on the underlying latent variable.

[140] Goodman (1997) discusses the difference between positive and negative behavioural measures.

Though the NLSCY behavioural scales differ from the SDQ subscales used in the MCS, there are several notable similarities between the results shown here in Table 5.13 and those for the non-cognitive measurement model presented in Table 4.13. In both tables, the hyperactivity measure has the highest factor loading in all periods, indicating that in both countries a high score on the hyperactivity scale will have a larger negative impact on the latent measure as compared to a high score in another other category.

There is variation in the loadings for the NLSCY behaviour measures across periods, which shows that although the questions used to generate the measure were consistent over time, the presence of certain behaviours has a larger influence on the generated factors during different stages of childhood. This variation in loadings corresponds with the changing distribution in the scores over time; a higher score during later periods is more indicative of a child being significantly different from the average.

Panel B of Table 5.13 provides the estimates for $\Phi^N_{j,t}$ in Equation 5.6. There are significant differences in how the non-cognitive indicators are measured for children of the same level of ability but of differing race or gender. When interpreting these coefficients, it ought to be noted that because the behaviours are based on parental perception of behaviour, the observed differences do not necessarily reflect different levels of behaviour, merely perceived differences. When it comes to gender, there are different perceptions of what is appropriate behaviour; an aggressive boy might be seen to be tough, while a girl behaving in an identical way might be perceived to be physically aggressive. As for race, there is evidence that parents from different ethnic backgrounds report the same behaviour differently. Studies from the US and the Netherlands show that when parent-reported behaviour is compared to reports from another observer, there are differences in reporting by ethnicity (Bevaart et al., 2012; Zwirs et al., 2006).

Bearing all these considerations in mind, it is possible to examine the coefficients for the co-variates. White children tend to have higher reported levels of physical aggression at all ages, higher reported conduct disorder symptoms for Cycles 3 and 5, as well as higher levels of emotional disorder during Cycle 5. The effect size of this measurement error for race is large, accounting for between a fifth and a half of a standard deviation increase in scores compared with these children's non-white peers. Across all ages, girls tend to have lower reported levels of hyperactive behaviour, emotional disorder-anxiety, physical aggression and property offences. From Cycle 4 onwards, girls have higher reported levels of indirect aggression. The magnitude of this difference in the perceived behaviour of boys compared with girls grows over time; by Cycle 6 when the children are ages 10–11, being female corresponds with having nearly a full standard deviation lower score on the hyperactivity measure and 20% of a standard deviation higher levels of indirect aggression. A child's age at the time of interview has small but significant effects on many of the reported measures, but these values are generally low.

Table 5.13 Measurement Model: Non-Cognitive Ability — Parameter Estimates

PANEL A: Latent Variables

	Cycle 3 — *Age 4/5*	Cycle 4 — Age 6/7	Cycle 5 — *Age 8/9*	Cycle 6 — *Age 10/11*
	θ_3^{NC}	θ_4^{NC}	θ_5^{NC}	θ_5^{NC}
Hyperactivity-Inattention	–1	–1	–1	–1
	–	–	–	–
Emotional Disorder-Anxiety	–0.484***	–0.412***	–0.52***	–0.664***
	(0.040)	(0.048)	(0.038)	(0.044)
Physical Aggression	–0.882***	–0.553***	-0.69***	–0.587***
	(0.071)	(0.041)	(0.047)	(0.040)
Indirect Aggression	–0.360***	–0.312***	–0.401***	–0.394***
	(0.048)	(0.032)	(0.039)	(0.035)
Property Offences	–0.680***		–0.493***	–0.368***
	(0.057)		(0.034)	(0.024)

PANEL B: Observed Covariates[1]

	Cycle 3 — *Age 4/5*			Cycle 4 — *Age 6/7*			Cycle 5 — *Age 8/9*			Cycle 6 — *Age 10/11*		
	White	Female	Age (mos.)	White	Female	Age (mos.)	White	Female	Age (mos.)	White	Female	Age (mos.)
Hyperactivity-Inattention	-0.082	–0.201***	–0.102***	-0.095	–0.242***	–0.038***	-0.130	–0.359***	–0.058***	–0.066	–0.363***	0.002
	(0.086)	(0.044)	(0.028)	(0.108)	(0.045)	(0.010)	(0.116)	(0.042)	(0.021)	(0.141)	(0.043)	(0.017)
Emotional Disorder-Anxiety	0.162	–0.101**	–0.066**	0.161	0.004	0.056***	0.264**	–0.034	–0.055***	0.149	–0.074	0.022
	(0.108)	(0.041)	(0.026)	(0.126)	(0.042)	(0.009)	(0.125)	(0.044)	(0.021)	(0.143)	(0.047)	(0.019)
Physical Aggression	0.341***	–0.263***	–0.039	0.302***	–0.316***	–0.019*	0.346***	–0.203***	-0.008	0.469***	–0.336***	–0.043***
	(0.076)	(0.039)	(0.029)	(0.103)	(0.041)	(0.010)	(0.121)	(0.044)	(0.025)	(0.088)	(0.043)	(0.016)
Indirect Aggression	0.104	0.045	-0.056*	0.057	0.171***	0.004	0.197*	0.171***	-0.010	0.078	0.211***	0.044***
	(0.122)	(0.048)	(0.032)	(0.130)	(0.047)	(0.011)	(0.102)	(0.046)	(0.025)	(0.138)	(0.048)	(0.016)
Property Offences	0.197***	–0.124***	–0.117***				0.215**	–0.270***	0.002	0.183	–0.260***	0.044***
	(0.075)	(0.043)	(0.020)				(0.102)	(0.041)	(0.024)	(0.132)	(0.041)	(0.015)

Notes:
[1] Parameter estimates for the observed covariates reported as the standard deviation change in the observed behaviour score associated with a unit change in the covariate.
– All models are estimated using provided survey weights and bootstrap weights.
– *p<0.1 , **p<0.05, ***p<0.01
– Standard errors in parentheses.

Measurement Model: Parental Investment

Determining the Factor Structure

The NLSCY contains a pre-validated set of scales based on the measures of parental investment included in the survey. Though the documentation provided with the NLSCY included an analysis of the underlying factor structure, Statistics Canada based this factor structure on data collected in Cycle 1 of the NLSCY across children ages 2–11. It was important for the present analysis to confirm that this structure held within the present data for all cycles used in the analysis. Thus, EFA was done separately for each cycle and the results are presented below in Table 5.14.

Table 5.14 Parental Investment Measures: EFA Measures of Fit

	Eigenvalue	RMSEA	CFI	TLI	SRMR
Age 4/5 - Cycle 3					
One Factor	4.961	0.073	0.597	0.543	0.131
Two Factor	2.271	0.057	0.784	0.720	0.095
Three Factor	1.976	0.039	0.916	0.874	0.055
Four Factor	1.228	0.033	0.947	0.907	0.044
Age 6/7 - Cycle 4					
One Factor	4.271	0.071	0.611	0.559	0.139
Two Factor	2.431	0.069	0.681	0.587	0.105
Three Factor	2.206	0.039	0.911	0.867	0.052
Four Factor	1.108	0.032	0.951	0.913	0.042
Age 8/9 - Cycle 5					
One Factor	4.842	0.052	0.758	0.726	0.132
Two Factor	2.388	0.061	0.718	0.635	0.087
Three Factor	1.871	0.037	0.908	0.862	0.052
Four Factor	1.180	0.033	0.939	0.892	0.038

Notes:

RMSEA = root mean square error of approximation. CFI = comparative fit index.
TLI = Tucker-Lewis index. SRMR = standardised root mean square residual.

Table 5.14 reports the relevant diagnostic measures from the EFA on the 18 measures of parental investment. These statistics indicate that there are three latent factors for parental investment across Cycles 3–6 of the NLSCY. In all three cycles, the eigenvalues are above one for four factors, but there is a substantial drop-off in the eigenvalue for the fourth factor compared to the third factor. For the three-factor model; the reported RMSEA is significantly under 0.05; CFI and TLI are above 0.90; and SRMR is below 0.08. As the goal of EFA is to identify the fewest possible factors required to model the data, the fit statistics point towards the use of a three-factor model.[141]

Table 5.15 provides the factor structure identified by the EFA for the three-factor model with 'X' marking the factor which had the largest significant factor loading for

[141] The factor structure was compared for both three and four-factor models and found that the four factor model yielded marginal increases in model fit. This increased fit came at the expense of a consistent factor structure over time as the four-factor model changed from one cycle to the next.

the given variable. With the exception of 'tell the child they are bad', all variables load onto one of the three latent factors. This variable — 'tell the child they are bad' — was excluded from the analysis because it did not load onto any of the latent factors, and including it as a lone-variable would require modifications to the theoretical model.

Table 5.15 Parental Investment Measures: Factor Structure

	Positive Interaction	Ineffective Parenting	Consistent Parenting
Praise child	X		
Five minutes focused attention	X		
Laugh together with child	X		
Get annoyed with child		X	
Tell child they are bad			
Do something special with child	X		
Play sports with child	X		
Proportion of talk: praise		X	
Proportion of talk: disapproval		X	
Make sure child follows commands			X
Punish child if breaks rules			X
Child gets away with breaking rules			X
Angry while punishing child		X	
Punishment depends on mood		X	
Problems managing child		X	
Child able to avoid punishment			X
Child ignores punishment			X
Need for repeated discipline		X	

Notes:
- To comply with Statistics Canada regulations EFA coefficients are not provided.
- EFA was conducted using provided bootstrap weights.

The factor structure remains relatively consistent across all cycles. The sole exception was the measure 'punishment depends on mood' which had the highest loadings on the second factor for Cycles 3 and 4, compared to the third factor for Cycle 5. This variable was included with the 'ineffective parenting' factor for all cycles. This provides consistency across cycles and aligns with the factor structure used by Statistics Canada.

The results from the EFA on the relevant subsample confirm the findings of Statistics Canada that the parenting scale is defined by three underlying factors of parental investment. This factor structure is consistent across Cycles 3–6 of the NLSCY and supports the theoretical constructs of parenting behaviours discussed in Chapter 2. For ease of interpretation, the names used to identify each factor in this paper match those given by Statistics Canada. The first factor is *positive interaction*, the second is *ineffective parenting* and the third is *consistent parenting*. Full numerical results from the EFA are not presented in this thesis as they have not been released from the RDC to comply with confidentiality requirements.

Results from the Measurement Model

The three factors identified above were included in the full empirical model. Table 5.16 presents the parameter estimates that correspond with $\alpha^s_{q,t}$ in the measurement model represented by Equation 5.7. Though the model as written in Equation 5.7 allows for covariates, no observed covariates were consistently correlated across the measurement model for parental investment, so Z^s_t is omitted from the equation for the analysis.

Table 5.16 Measurement Model: Parental Investment - Parameter Estimates

	Cycle 3 *Age 4/5*			Cycle 4 *Age 6/7*			Cycle 5 *Age 8/9*		
	I^1_3	I^2_3	I^3_3	I^1_4	I^2_4	I^3_4	I^1_5	I^2_5	I^3_5
Positive Interaction									
Praise child	1.000			1.000			1.000		
	–			–			–		
Five minutes focused attention	1.121			1.138			1.078		
	(0.046)			(0.058)			(0.039)		
Laugh together with child	1.043			1.131			1.163		
	(0.046)			(0.059)			(0.042)		
Do something special with child	0.788			0.886			0.889		
	(0.045)			(0.059)			(0.048)		
Play sports with child	0.851			0.895			0.817		
	(0.051)			(0.061)			(0.040)		
Ineffective Parenting									
Get annoyed with child		1.000			1.000			1.000	
		–			–			–	
Proportion of talk: praise		0.958			0.945			0.884	
		(0.066)			(0.053)			(0.051)	
Proportion of talk: disapproval		1.178			1.072			0.989	
		(0.064)			(0.051)			(0.057)	
Angry while punishing child		0.987			0.954			1.114	
		(0.059)			(0.057)			(0.050)	
Punishment depends on mood		0.869			0.724			0.888	
		(0.065)			(0.056)			(0.046)	
Problems managing child		1.307			1.156			1.174	
		(0.064)			(0.064)			(0.051)	
Need for repeated discipline		1.146			1.100			1.164	
		(0.056)			(0.057)			(0.047)	
Consistent Parenting									
Make sure child follows commands			1.000			1.000			1.000
			–			–			–
Punish child if breaks rules			1.040			1.259			1.086
			(0.061)			(0.073)			(0.085)
Child gets away with breaking rules			1.332			1.457			1.507
			(0.094)			(0.103)			(0.084)
Child able to avoid punishment			1.137			1.494			1.429
			(0.082)			(0.124)			(0.088)
Child ignores punishment			1.185			1.464			1.390
			(0.102)			(0.123)			(0.082)

Notes:

– $p<0.01$ for all variables.

– Standard errors shown in parentheses.

Since the factor structure for the parenting scales is consistent over time, it is possible to compare the magnitude of factor loadings at different ages. If the factor loadings vary, it indicates that the relative information provided by each measure changes over time. Change in the information provided by a measure would likely occur if some measures were to be more widely distributed at one point compared with another. In general, there are no drastic changes in the relative factor loadings over time. Slight variation is observed in all three of the measures, with a few notable differences.

The first latent variable for parenting is constructed using measures that correspond with positive interactions between the parent and child. All five measurements appear to provide similar levels of information about the underlying latent factor with loadings ranging from 0.788 and 1.163. The relative ranking of the factor loadings remains mostly stable from Cycle 3 to Cycle 5, with only minor changes in the overall ranking of magnitudes. For example, in Cycle 3, the loading on 'five minutes of focused attention' is 0.078 higher than on 'laugh together with the child'; however, by Cycle 5 the difference has reversed, with 'laugh together with the child' being 0.085 higher. These changes in magnitude are all relatively small, with no indicator moving more than one position in the relative ranking. This reveals that the amount of information about the underlying construct contained in each indicator remains stable over time.

The second latent variable is constructed using measures that correspond with ineffective parenting. Across all three cycles the highest factor loading is on the measurement of 'problems managing the child'. In Cycle 3, this measure had a loading that was 130.7% of the loading on 'get annoyed with child.' This indicates that variability in the frequency that parents have 'problems managing' their child is able to capture more of the underlying construct of ineffective parenting. In Cycle 5, the spread of the factor loadings has decreased; this means that all of the indicators provide similar amounts of information about the underlying factor and likely captures the increased spread seen in these measures during Cycle 5.

The third latent variable is defined using measures of consistent parenting. For this latent factor, the loadings on the last three measures are substantially higher than those on the first two. The loadings on 'child gets away with breaking rules', 'child is able to avoid punishment' and 'child ignores punishment' are approximately 150% the magnitude of those on 'make sure the child follows commands' and 'child is punished if they break rules'. This means that the last three indicators contain more information about the underlying latent variable of 'consistent parenting'.

Structural Model:

Table 5.17 presents the parameter estimates for the structural model. These estimates in Panel A correspond with the coefficients Γ_t and B_t in Equation 5.1. Panel B contains the estimates of the parameter Λ_t which measures the impact of family income, a time varying covariate which corresponds with X_t in Equation 5.1.

Table 5.17 Structural Model: Parameter Estimates

PANEL A: Latent Factors[1]

	Cognitive Ability (θ^C_{t+1})			Non-Cognitive Ability (θ^N_{t+1})		
	θ^C_4	θ^C_5	θ^C_6	θ^N_4	θ^N_5	θ^N_6
Lagged Ability:						
Cognitive (θ^C_t)	0.457***	0.360***	0.456***	0.122***	-0.067**	-0.032
	(0.047)	(0.033)	(0.036)	(0.036)	(0.028)	(0.027)
Non-Cognitive (θ^N_t)	0.102***	0.044	0.045*	0.743***	0.880***	0.866***
	(0.024)	(0.029)	(0.023)	(0.028)	(0.024)	(0.019)
Lagged Parental Inputs:						
Positive Interaction (I^1_t)	0.080***	0.047*	0.040*	0.190***	0.225***	0.214***
	(0.022)	(0.024)	(0.021)	(0.031)	(0.025)	(0.028)
Ineffective Parenting (I^2_t)	-0.111***	-0.014	-0.057***	-0.539***	-0.481***	-0.521***
	(0.023)	(0.018)	(0.020)	(0.025)	(0.022)	(0.028)
Consistent Parenting (I^3_t)	0.035	0.055**	0.106***	0.202***	0.273***	0.280***
	(0.026)	(0.023)	(0.022)	(0.027)	(0.026)	(0.024)

PANEL B: Observed Covariates[2]

	Cognitive Ability (θ^C_{t+1})			Non-Cognitive Ability (θ^N_{t+1})		
	θ^C_4	θ^C_5	θ^C_6	θ^N_4	θ^N_5	θ^N_6
Second Income Quintile	0.285***	0.310***	0.190**	-0.015	0.083	0.256***
	(0.095)	(0.089)	(0.075)	(0.122)	(0.104)	(0.097)
Third Income Quintile	0.408***	0.402***	0.018	-0.033	-0.061	0.104
	(0.110)	(0.118)	(0.090)	(0.135)	(0.105)	(0.113)
Fourth Income Quintile	0.402***	0.447***	-0.006	-0.111	0.123	0.071
	(0.117)	(0.123)	(0.095)	(0.166)	(0.132)	(0.134)
Fifth Income Quintile	0.448***	0.325***	0.168	0.115	0.138	0.390***
	(0.126)	(0.116)	(0.105)	(0.165)	(0.135)	(0.132)

Notes:
[1] Parameter estimates for the latent factors reported as the standard deviation change in the latent score associated with a standard deviation change in the lagged latent score.
[2] Parameter estimates for the observed covariates reported as the standard deviation change in the latent score associated with a unit change in the covariate.
⁻ All models are estimated using provided survey weights and bootstrap weights.
⁻ *p<0.1 , **p<0.05, ***p<0.01
⁻ Standard errors in parentheses.

As in the previous empirical application, both cognitive and non-cognitive ability are self-productive, with higher levels in one period predicting higher levels of the same type of skill in the next period. For both types of skill, this autoregressive effect is statistically significant over all three periods of measurement. The magnitude of the estimated self-productivity is much larger for non-cognitive ability. One explanation is that non-cognitive ability is measured using parental reports, which suffer from parent reporting bias. This bias is likely consistent across periods, and partially captured by the model; it is not seen in the cognitive measures which are the result of objective cognitive tests administered by the NLSCY.

For cognitive ability, there is strong evidence of cross-productivity of non-cognitive ability in Cycles 4, while neither Cycles 5 nor 6 are statistically significant at the 5% level. The magnitude of this cross-productivity appears to decrease over time with the effect being 0.102 in Cycle 4 but 0.045, and only slightly significant, for Cycle 6. This indicates that having a composite non-cognitive score one standard deviation above the mean in Cycle 3 corresponds with being 0.102 standard deviations above the mean for the cognitive ability in Cycle 4. For non-cognitive development, there is also cross-productivity for the first two periods of development, but this cross productivity falls in magnitude and is no longer statistically significant for the final period.

The three latent variables for parental investment are all significant determinants of non-cognitive skills during every period measured in the present model but are only significant at certain points for cognitive skills. These results indicate that the underlying parenting constructs used in this model are critical to all periods of non-cognitive development, but not all periods of cognitive ability. By closely examining the significance and/or magnitude of the coefficient over time, it is possible to see when these specific types of investment are most effective in non-cognitive and cognitive skill development.

For positive interaction, which was constructed using measures of time spent in parent-child activities, the coefficient is a significant predictor of cognitive ability across all 3 periods, though only at the 10% level for the second and third periods. The magnitude of this coefficient falls over time, indicating that the impact of such parenting behaviours is largest in earlier stages of development. This factor is significant at the 1% level for non-cognitive ability at all ages, with the magnitude of the coefficient remaining relatively stable over time. The coefficient for non-cognitive ability is highest in the second period of the model with one standard deviation increase in the parenting construct corresponding with an 0.225 higher score on non-cognitive ability.

The ineffective parenting factor, which was constructed using measures of the role that a parent's mood played in shaping their interactions with the child, is statistically significant for the cognitive ability at age 5–6, and again at ages 10–11, but not for the intermediate period. We may extrapolate from this that children's cognitive ability is sensitive to ineffective parenting in pre-primary years, and later in primary school but that there is no observable increase in cognitive ability from this type of parental investment during the early primary years. However, the ineffective parenting factor is also significant for non-cognitive ability at all ages, which indicates that a parent's mood and interactions produce consequences that have benefits beyond cognitive skills. Ineffective parenting appears to be especially detrimental, to development with the coefficients ranging from -0.481 to -0.539, meaning that parents scoring one standard deviation higher on this construct have children whose non-cognitive ability is approximately half a standard deviation below the mean. This follows a similar pattern to the similar construct of *parent-child interactions* in the UK data, which also remained significant across all periods with a the highest coefficient seen at age 7. For both Canada and the UK, it is possible that this coefficient captures some reverse causality, whereby parents who have children with poor behaviour respond by minimising their interactions.

The consistent parenting factor is significant for the final two periods for cognitive ability and across all periods for non-cognitive ability. This factor was constructed using questions about discipline and enforcement of rules. For both types of ability, the magnitude of the coefficient on this factor grows over time, indicating that consistency in discipline plays the largest role in the skills of older children.

Panel B of Table 5.17 presents the parameter estimates for the parameter Λ_t in Equation 5.1 of the structural model. Though these estimates of the effect of X_t (family income) can be thought of as control variables, they are critical to demonstrating the importance of including them separately from parental investment in the model of skill formation. The coefficients presented in Table 5.17 show that income is a significant predictor of higher cognitive ability in Cycles 4 and 5. Being in any quintile above the lowest corresponds with scores at least 0.285 standard deviations above the mean. These coefficients are even larger than those found in the previous empirical application and three to four times the magnitude of those on positive interaction and consistent parenting. Therefore, any model including family income with investment would likely vastly overstate the effect of parental behaviour. For non-cognitive ability, the Canadian data appears to show little effect of income, with the only significant coefficients being in Cycle 6. This differs from the UK data which found a small but consistent relationship over all periods.

Table 5.18 presents the parameter estimates for the covariates for initial cognitive and non-cognitive ability. Unsurprisingly, being in the top two income quintiles corresponds with slightly higher scores on the initial measures of cognitive ability, as does increasing maternal age and maternal education. Contrary to expectations, low birth weight does not significantly predict cognitive scores. The effect of education is large, with children whose parents hold a higher degree having cognitive scores at age 3 that are 0.495 standard deviations above the mean.

Table 5.18 Structural Model: Parameter Estimates Initial Period Covariates

	Cognitive Ability (θ_0^C)	Non-Cognitive Ability (θ_0^N)
Second Income Quintile	–0.065	–0.069
	(0.076)	(0.114)
Third Income Quintile	0.151	–0.061
	(0.094)	(0.128)
Fourth Income Quintile	0.265***	–0.123
	(0.092)	(0.135)
Fifth Income Quintile	0.300***	-0.325*
	(0.103)	(0.173)
Low Birth Weight	0.057	0.242*
	(0.130)	(0.138)
Mother's Age at Birth	0.033***	0.049***
	(0.006)	(0.006)
Highschool Graduate	0.124	–0.135
	(0.080)	(0.108)
Some Post-Secondary	0.197***	–0.130
	(0.068)	(0.115)
Trade School or College Diploma	0.235***	–0.158
	(0.072)	(0.104)
Bachelor's Degree	0.413***	–0.125
	(0.087)	(0.127)
Higher Degree	0.495***	–0.702***
	(0.134)	(0.188)

Notes:
- All models are estimated using provided survey weights and bootstrap weights.
- *p<0.1 , **p<0.05, ***p<0.01
- Parameter estimates for the observed covariates reported as the standard deviation change in the latent ability scale associated with a unit change in the covariate.
- Standard errors in parentheses.

For early non-cognitive ability most of the covariates do not show significant effects. Parental education is only significant for the highest level, indicating that compared to children whose parents hold no qualification, those holding a higher degree have non-cognitive scores that are 0.702 standard deviations *lower*. As discussed earlier in this chapter, the parent perceived frequency of certain behaviours is likely to suffer from some bias. This finding indicates that parents who have a higher degree might be more critical of their children's behaviour. The coefficient on low birthweight is only significant at the 10% level and shows that children born below 2500 grams are 0.242 standard deviations below their normal weight peers.

Table 5.19 provides the estimated coefficients for ϕ_t^I in Equation 5.2. Family size is only statistically significant for positive interaction. Furthermore, with the effect growing between Cycle 3 and 4 before disappearing in Cycle 5. The results show that during Cycle 3, when the children are ages 4–5, each additional sibling in a family corresponds with a 0.182 standard deviation decrease in positive interaction. Furthermore, during Cycle 4 when the children are aged 6–7 ,each sibling corresponds with a 0.245 standard deviation increase in positive parenting. Being a single parent appears to have mixed implications for the 'ineffective parenting' input, with single parent status increasing the amount of ineffective parenting at Cycle 3, but decreasing it at Cycle 5. For consistent parenting, the impact is more uniform with both Cycles 3 and 4 showing single parents having significantly lower levels of this behaviour.

Table 5.19 Structural Model: Parameter Estimates Covariates for Investment

	Cycle 3		Cycle 4		Cycle 5	
	No. of Siblings	Single Parent	No. of Siblings	Single Parent	No. of Siblings	Single Parent
Positive Interaction (I_t^1)	-0.182**	-0.165	-0.245***	0.079	0.044	0.047
	(0.072)	(0.133)	(0.093)	(0.131)	(0.101)	(0.131)
Ineffective Parenting (I_t^2)	0.103	0.392***	0.086	-0.113	-0.058	-0.278**
	(0.078)	(0.135)	(0.092)	(0.129)	(0.089)	(0.116)
Consistent Parenting (I_t^3)	-0.069	-0.268**	0.016	-0.321**	-0.017	0.043
	(0.088)	(0.130)	(0.116)	(0.147)	(0.100)	(0.161)

Notes:
[1] Parameter estimates for the observed covariates reported as the standard deviation change in the latent parenting scale associated with a unit change in the covariate.
– All models are estimated using provided survey weights and bootstrap weights.
– *p<0.1 , **p<0.05, ***p<0.01
– Standard errors in parentheses.

Taken together, the results above show that there are specific sensitive periods for these types of parental investment and that cognitive and non-cognitive ability show signs of cross- and self-productivity. Unlike past research, which found that cognitive ability demonstrates less sensitivity to investment as children age, this model shows that different types of parenting have varying periods of sensitivity for cognitive ability. This ability to differentiate between types of skill is a feature not captured by the original work of Cunha and Heckman (2008). Even though an analysis of data from the UK, by Hernández-Alava and Popli (2017), differentiated between several categories of parent-child activity, to my knowledge, this study is the first to differentiate between types of parental investment in Canadian data.

5.5 Discussion

This chapter provides the second empirical application of the modified model of skill formation introduced by this thesis. This analysis uses longitudinal data from the NLSCY to estimate the model within a Canadian context and provides further support for using the empirical model to capture skill development in a variety of contexts. This proof of concept is in addition to the valuable findings of the empirical estimates of the model. These findings can be used to create policy recommendations for the effects of various types of parenting, and to help policymakers identify the periods of childhood when they are especially beneficial for children.

This specification of the model highlights the ability of this modified skill formation model to be applied to a variety of different types of parenting constructs. More specifically, while the UK analysis presented in Chapter 4 used the model to differentiate between various parent-child activities, this empirical application uses parental investment measures that focus on underlying general parenting practices. This allows me to provide the first estimates which use the skill formation model to measure how parenting constructs directly influence cognitive and non-cognitive skill development.

Building on the existing knowledge on skill formation, the results of my analysis show that cognitive ability and non-cognitive ability are both strongly persistent over time. I find that early non-cognitive skills are strong determinants of cognitive ability, though this cross-productivity diminishes over time. The cross-productive effect of cognitive ability on non-cognitive ability is significant over all periods, but also decreases in magnitude over time, indicating that non-cognitive skills are reliant on the presence of cognitive ability for their growth at all stages of development. As with the MCS data, this persistent influence of parenting on non-cognitive skills points to the importance of interventions for children with early deficits in non-cognitive ability.

The evidence of both cross-productivity and self-productivity implies that children who face deficits in either type of skill early in life will benefit less from future investment and are likely to fall further behind. This points to the need for strong early intervention programs which identify the means to reduce these early childhood gaps. From my own analysis, the most influential parental input is 'ineffective parenting': a parent who is one standard deviation below the mean on this factor, has, on average, a child approximately a half a standard deviation below his or her peers on the latent construct of non-cognitive ability. If it is possible to identify specific pathways to promote more effective parenting styles this could lead to vastly improved outcomes for a subset of Canadian children.

By adding exploratory factor analysis as an initial step in modelling parental investment, this model allows for parenting behaviours that might not all load onto a singular construct. Unlike the first empirical chapter, this Canadian application focused on general parenting attitudes instead of on specific types of parent-child interaction. The model explicitly separated measures of family resources from parental behaviour as the goal was to find parenting behaviours that can be encouraged across the SES spectrum. Again, the model showed that it is possible to find significant effects of parenting — separate from the effect of family resources — with the coefficients on the parenting constructs being similar or larger in magnitude to the effect of family income alone.

As in the UK analysis, the Canadian model showed the capacity of the model to provide a more nuanced understanding of critical and sensitive periods of parental investment. Having distinguished between types of parenting behaviours, I find that the previous critical periods may overgeneralise and overlook certain specific behaviours. Cunha and Heckman (2008) and Cunha et al. (2010) both emphasised the sensitivity of cognitive ability to early investment and non-cognitive ability to investment at later ages. The findings from this analysis show that there is a more complex explanation, with specific types of parenting behaviour showing benefits at all stages. This finding presents a more optimistic outlook than the initial findings which appeared to show little-to-no impact of parenting on later cognitive skill development.

As this is the first analysis that includes parenting styles within the skill formation models, it is only able to scratch the surface of the dynamics of skill formation within this context. Nevertheless, the results from this analysis clearly highlight the applicability of the model, along with the consistency of the findings concerning the cross-productivity of skills, the effect of certain covariates and the strong role played by family income. With these findings it is possible to begin the process of further research; potential projects include exploring other data that contains similar measures and experimental interventions, such as the Fast Track program in the United States. Although the results from this data will also be correlational, confirming the results from this study using other data will provide further evidence to suggest causality, as well as alternative ways of interpreting the results from my study.

CHAPTER 6

Discussion

This chapter serves to conclude my thesis by consolidating the theoretical framework introduced in Chapter 3 with the empirical findings presented in Chapter 4 and Chapter 5. To accomplish this, the chapter is organised in the following manner. First, Section 6.1 summarises the key findings and revisits the research questions. Drawing on these findings, the methodological and empirical contributions of this thesis are presented in Section 6.2. Next, Section 6.3 explains the implications that my empirical findings have for public policy. I end the chapter by considering the limitations of this thesis and presenting recommendations for future research in Section 6.4 and Section 6.5 respectively.

6.1 Summary of Findings

At the beginning of this thesis, I introduced three research questions. Based on these questions, I set out to define a theoretical framework to explain the role that parents play in childhood skill development and to apply this proposed framework to existing data in order to obtain policy-relevant empirical estimates. In this section, I return to the research questions and examine how they are addressed by the theoretical framework introduced in Chapter 3 and the empirical findings presented in Chapter 4 and Chapter 5. Before explaining how this thesis responds to each individual research question, I present a brief summary of the empirical findings.

General Overview

In Chapter 3, I defined a theoretical framework to model the developmental trajectories of cognitive and non-cognitive skills in primary school children and to capture the role that parenting plays in the joint evolution of these skills. This framework is an updated version of the economic model of skill formation originally presented by Cunha and Heckman (2007). I proposed modifying their model to distinguish between financial investments in children and other parental inputs to development. With this goal, I introduced both data-driven and theoretical approaches that build on the literature from psychology and education to aid in the identification of such parental inputs.

Chapter 4 and Chapter 5 presented applications of my modified skill formation model which examined data from longitudinal cohort studies in the UK and Canada, respectively. In both countries, three parenting inputs are identified. Each of these inputs has a significant effect on the development of both cognitive and non-cognitive skills, with differing periods of sensitivity to each type of parental input.

The first empirical study analysed a sample of 8,379 children from the UK Millennium Cohort Study (MCS). Using repeated measures of parenting behaviors at ages 3, 5, and 7, three separate latent parental inputs are identified: *literacy activities*, *parent-child interactions* and *academic activities*. I find that both *literacy activities* and *parent-child interactions* have sizeable effects on early childhood cognitive ability, but this decreases over time. For non-cognitive ability, all three inputs have significant effects, but again the marginal rate of return decreases over time. As well as identifying latent parenting factors, the model estimates also show that cognitive ability and non-cognitive ability are both strongly persistent over time. Furthermore, early non-cognitive skills are strong determinants of cognitive ability, indicating that there is a significant cross-productive effect of non-cognitive ability in the early formation of cognitive ability.

The second empirical application of the model estimated the model in a Canadian context using a sample of 1,234 children from the National Longitudinal Survey of Children and Youth (NLSCY). Again, three separate latent parental inputs were identified. These capture different aspects of parenting style and are labelled *positive interaction*, *ineffective parenting* and *consistent parenting*. The results from my analysis show that all three types of parental input are correlated with non-cognitive ability and the magnitude of these effects remains largely consistent over time. For cognitive ability, the results are more complex. They show that cognitive ability is sensitive to *ineffective parenting* and *consistent parenting* until the final period of the study, when the children are aged 10–11. At first glance, this continued sensitivity of cognitive ability to parental input appears contrary to the findings of Cunha and Heckman (2008) and Cunha et al. (2010) which showed that cognitive ability had limited sensitivity to investment at during later periods of childhood. A feasible explanation is that the findings from my model allow for a more nuanced measure of sensitivity when compared to previous studies. When all the measures of investment are collapsed into a single factor, the small but significant effects from one type of input might be outweighed by the lack of significance from another input. As a result, the sensitive periods I identify are missed by the consolidated measure of investment used by Cunha and Heckman. Separating the types of investment shows that specific types of parenting behaviour influence development at each stage.

Revisiting the Research Questions

Below, I discuss how the chapters summarised above serve to address the research questions which originally motivated this thesis.

1. **How can the literature from psychology and economics be consolidated within the field of education to form a theoretical framework to explain parents influence the development of cognitive and non-cognitive skills?**

To address this question, I identified seven key findings from the relevant literature at the end of Chapter 2 and then used these findings to help contextualise the theoretical framework which I presented in Chapter 3. For the sake of brevity, the seven findings are not listed again below, but they can be thought of as addressing two main areas of the literature: empirical and methodological.

Empirical Literature: There is significant overlap between the empirical research in the fields of education, economics and psychology. Though researchers in each of these fields reference the findings from other disciplines, some of the key parallels between these strands of research are obscured by minor differences in research terminology.

Methodological Literature: Neither economics nor psychology provides the full set of methodological tools required to model and estimate skill development. While economics provides a mathematical framework to model how family characteristics, household resources and parental behaviours shape the outcomes of children, this mathematical model is driven by theories of scarcity. If we were to assume, as dictated by economic theory, that parental behaviour is determined by the financial and time costs faced by parents, these models are unable to differentiate between inputs that require the same expenditure of time or money by parents, but result in drastically different outcomes. Additionally, even though the inputs included in these types of models are often based on psychometric constructs, economists do not necessarily use psychological theories to help specify their specific model inputs. On the other hand, while psychological research pays greater attention to understanding the mechanisms which explain various constructs and measures, empirical studies from psychology do not have the same well developed, recursive framework that is present in human capital models.

At the end of Chapter 2, I argue that the field of education provides the context for building on a variety of theoretical perspectives and empirical findings to develop a better understanding of skill development in children. Within the multidisciplinary lens of education research, Chapter 3 presents my own model in which I have adapted the mathematical framework from economic models of human capital so that I can accommodate theoretical constructs and psychometric measures taken from the field of psychology. This provides a theoretical framework to help identify different types of parental inputs and then to explain how these might uniquely influence development.

2. Can existing models of skill development be improved in order to distinguish between parental behaviours and socio-economic resources?

To address the second research question, I have proposed an approach for identifying inputs to development that combines data-driven exploratory factor analysis with theoretical justification from the literature discussed in Chapter 2. While the specific details of this methodology are provided in Section 3.3, the empirical applications presented in Chapter 4 and Chapter 5 demonstrate how this methodological approach provides more nuanced estimates of the role that parental input plays in skill formation. These precise estimates highlight that the modification is indeed an improvement on existing models of skill development. The importance of distinguishing between different types of parental inputs will be revisited in Section 6.2.

3. What can empirical estimates from a variety of contexts tell us about the role that specific parenting constructs play in the trajectory of childhood skill formation?

The findings from the two empirical chapters provide a wealth of insight into the development of cognitive and non-cognitive skills in primary school children. The brief summaries at the beginning of this section highlight some of the key findings, but for the full response to this research question the reader is directed to the discussions of the empirical estimates provided at the ends of Chapter 4 and Chapter 5.

6.2 Contributions to the Field

Methodological Contributions

The framework that I introduced in Chapter 3 makes two main methodological contributions to research on parenting and skill development.

My first methodological contribution is to define a model of skill formation which allows for multiple types of parental investment and specifically differentiates between the effects of parental behaviour and family socio-economic resources. The literature in Section 2.2 outlines the many ways in which SES can shape a child's development, and my own empirical findings demonstrate how separating SES from other types of parental inputs allows for a better understanding of the role of parenting behaviours.

To define the various types of parental investment, I proposed an approach which begins with data-driven EFA to identify the types of parenting from the data, and then situates these findings within existing theoretical constructs. For readers with a background in psychology or education, applying EFA in this way is a common approach to interpreting data, but for readers grounded in economics, it is more typical to use existing constructs when examining the data, and only use PCA as a form of data reduction. In this thesis, I have chosen an approach which uses EFA to identify parental inputs, but then applies them within a model strongly rooted in economics.

The identification of parental inputs is one example of the second major methodological contribution of this thesis. Namely, the consolidation of the theoretical and empirical findings from economics and psychology, under the multidisciplinary lens of education. By building on the strengths of each field, I am able to define a model which defines parental inputs to development more precisely than pre-existing models, thereby providing a better understanding of the trajectories of skill development in childhood.

Research which uses empirical findings from one field to define the measures used within the theoretical approaches from another field is not unique. However, in order to capitalise on the knowledge from psychology, education and economics, I propose that research must use theoretical justifications from each discipline when defining a model of skill formation. This differs from the existing research, which tends to use the theoretical lens of one discipline while drawing on the other for measurement tools or empirical strategies. In economics, this takes the form of human capital models which include pre-existing psychological measures as inputs or outputs in the model while relying on economic theories for the theoretical justification of the model. By contrast, in psychology, studies will apply economic strategies to aid in data reduction while using psychological constructs to frame the theoretical approach. Therefore, I have presented a methodology that draws on theoretical justifications from both economics and psychology and situates this combined approach within the field of education.

Empirical Contributions

In addition to confirming the methodological legitimacy of the proposed model, applying my empirical framework to data from Canada and the UK serves to provide new empirical estimates of skill formation in each country. These estimates make an important contribution to the literature on the role that parents play in skill development.

Earlier in this chapter, in Section 6.1, I reviewed the key findings from each of the two empirical applications. The findings add to the empirical knowledge about skill development in each country as well as identifying global patterns in the role that parents play in childhood skill formation.

In both the Canadian and UK data, I have found that childhood skill development and parental behaviour are strongly connected – regardless of family socio-economic circumstances. My modelling strategy has shown that specific parental inputs matter at different ages of a child's development. This empirical evidence supports the need for distinguishing between types of parental inputs to skill development and forms the basis of the policy implications that I will discuss in Section 6.3.

In the UK, I was able to examine the role of various types of interactions between parents and their children. This extended on the findings of Bono et al. (2016) and Hernández-Alava and Popli (2017), who use the MCS to find that parent-child interactions predict both cognitive and non-cognitive development at ages 3, 5 and 7. As my analysis used the same data as these studies, it is possible to use their findings as a point of comparison for the more specific estimates offered by my modelling strategy.

For my Canadian analysis, the data contained in the NLSCY led to an alternative definition of parental investment, by examining repercussions of parenting styles on children's outcomes. To my knowledge, a recursive economic model of skill formation has not been estimated using this data, but various studies have used the NLSCY data to examine the effect that these parenting styles have on behaviour and cognitive development. For example, Dooley and Stewart (2007) find that parenting style is highly predictive of the behavioural outcomes which I define as non-cognitive ability, and Baker and Milligan (2016) find that parenting practices are predictive of primary school measures of cognitive ability. The findings from my larger model confirm these existing estimates and help to build on the understanding of how skills develop in Canadian children.

6.3 Implications for Policy and Practice

Taken together, the results from the two empirical applications provide general findings which can help to shape the way that policymakers, parents and educational practitioners think about child development. The specific findings from each context also have policy implications that are relevant for each country. These country-specific recommendations will vary based on the existing structures, redistributive policies and education systems in each context. In countries where there is greater social and economic inequality, it is likely that separating SES from parenting behaviours will be more revealing, since prior analyses in these countries will have falsely attributed socioeconomic gaps in ability to the effect of parenting behaviours. In countries where extensive redistributive policies exist, socioeconomic gaps in cognitive and non-cognitive skills are narrower (Bradbury et al., 2015a). Therefore, models which combine SES with other parental inputs will be less likely to be biased by SES driven differences and more likely to capture the effect of parenting behaviours.

Regardless of the context in which the estimates are taken, the findings from my model can be used in several ways. First, they can serve as guidance for existing interventions which target the parents of at-risk youth. For example, the *Fast-Track Program* in the US has identified children at risk of poor academic, behavioural and emotional outcomes and has subsequently provided a comprehensive set of interventions which include parenting classes. Since the results from my model show that time spent in direct interaction with the child improves both cognitive and non-cognitive outcomes in primary school, it is logical to encourage parents to engage in these behaviours.

Although my analysis focused on activities that took place primarily outside the formal education system, child development is the product of multiple contexts including formal education. My findings can help practitioners to understand the role parenting behaviours play in explaining the gaps in cognitive and non-cognitive achievement that they witness in the children they teach. Moreover, there is some capacity for educational policy to find ways to encourage these behaviours both during and outside of school hours. For example, given that my findings demonstrate a strong link between non-academic parenting behaviours and early cognitive ability, it might be that, in contexts where parental engagement with schools exist, the school could find ways to encourage parents to participate in these activities. Naturally, the specifics of such interventions are far more complex than I have described, but my findings can serve as a starting point for developing evidence-based policies. Crucially however, my models provide policymakers and practitioners with essential indicative evidence of the potential of interventions of this nature for the long term development of children's abilities.

Finally, in addition to policy-based interventions, the findings from this thesis are also of direct interest to parents. Parents often want to learn more about the ways in which they can ensure the best outcomes for their children. The findings from my research can easily be synthesised to provide precise recommendations to improve the effectiveness of the time they spend with their children. My work may also change the discourse around children's development which has rightly tended to focus on the negative impact of low SES on children and not provided poorer parents with any indication of the agency they may have in addressing their children's needs.

6.4 Strengths and Limitations

Methodological Strengths and Limitations

As with any methodology that relies on observational data, my model is unable to conclusively prove a causal link between parental behaviours and childhood outcomes. Such proof is only possible with the use of experimental or quasi-experimental approaches, but such approaches are rarely possible on the nationally representative scales I have analysed. Even for smaller-scale studies, there are practical and ethical limitations that make it difficult to implement experiments which change the behaviours of some parents while controlling for other factors related to development. Fortunately, the recursive model I propose includes substantial control variables and can be applied to large data sets. For now, my proposed approach, using rich data from cohort studies, provides very robust estimates — the closest an analysis can get to identifying causality.

There is one small concern with the empirical model, and this is the possibility of reverse causality between childhood outcomes and parenting behaviours. More specifically, parents will respond to their children's behaviour and alter their own in response. This could take the form of increasing their investment if they believe children are struggling, or alternatively, reducing interaction if their children demonstrate high levels of negative behaviour and are therefore unpleasant to be around. In the next section of this chapter, I discuss ways to examine reverse causality and study this issue further.

Strengths and Limitations of the Data

My analyses made use of existing, longitudinal data. In Chapter 3, I explained how large sample sizes provide sufficient statistical power for the complexity of my modelling approach. Both the MCS and the NLSCY contained sufficiently large samples. The sample used for the UK analysis (n=8,379) provides especially robust results, while the Canadian analysis (n=1,234) is only slightly above the minimum sample size of 1,000 for this type of analysis. Although the smaller Canadian sample might reduce the number of statistically significant estimates, the results should still be considered quite robust.

In addition to the large sample sizes, another strength of the data used in this thesis is the wealth of available measures. Both the MCS and the NLSCY contain robust measures with high quality data collection. The wealth of measures provided by these studies allows me to include a variety of control variables in my analyses.

While the cohort studies have many strengths, any analysis which relies on secondary data will be limited by the need to depend on the predetermined measures which are included in each survey. In each empirical chapter, I discussed the specific limitations of the measures contained in each data set and explained how the choice of parental input measures was driven, in part, by the variables available in each survey.

6.5 Directions for Future Research

Future Applications of my Methodological Approach

My proposed framework can be applied to a variety of contexts and types of parental input. Each of these settings and types of input is a potential avenue for future research. There are several possible extensions that can be made to the analyses I have presented. Additionally, the model can be applied to data from other contexts in order to build the international body of evidence on skill formation and to provide further context within which to examine the results from the Canadian and British analyses.

Although the study presented in Chapter 4 is already an extension of existing research which estimates childhood human capital production functions in Britain, there are several further extensions that can be conducted using the MCS data. First, the analysis can be extended to capture the most recent sweep of the MCS (Sweep 6) which has only recently been released. Data from this sweep will allow me to see how development progress into the pre-teen years, and to examine whether or not the gaps developed in early childhood continue to persist in adulthood. Next, the analysis can be broadened to include measures of paternal input. In this thesis, the analysis was limited to maternal inputs only. This decision was made to maximise sample size, as the fathers' response rate on the MCS was substantially lower than the overall response. Nevertheless, extending the analysis to include paternal inputs is possible and would offer valuable insight into the precise roles that fathers play in child development. Including paternal inputs would allow us to understand if a father's behaviour can serve as substitutes (or complements) to maternal investment. Finally, I propose that future research applying my theoretical framework to the MCS ought to involve other measures of parental investment that are covered by the survey. For example, the studies by Hernández-Alava and Popli (2017) and Bono et al. (2016) both suggest alternative ways of defining investment. These types of investment include measures of the time spent in parent-child interaction and questions relating to parental attitude — including such measures may be useful in understanding the findings from my model.

Similarly, there are many options for extending my model in the Canadian data. As with the MCS, it is possible to re-examine the NLSCY using other measures of parental behaviour or to modify the analysis to examine the role of fathers. Furthermore, the data contained in the NLSCY would support the application of my model to older children, facilitating the estimation of how parental investment shapes development in adolescence and early adulthood. While the MCS followed only one age group over time, the NLSCY identified children from 0–11 years old in 1994 and followed them for the next 16 years. Studying how skills develop as children progress through secondary education would provide a more complete picture of the nature of human capital formation.

Building on my Empirical Findings

The findings from my empirical studies can be used to inform other types of research. Once the key periods in which cognitive ability is most sensitive to parental inputs have been identified, the next step is to design experimental studies which involve interventions to promote these types of parenting behaviours and then measure the children's outcomes. It might seem drastic to propose a programme which seeks to alter the way parents raise their children, but since the goal of the intervention would be to promote the behaviours captured by the parental inputs I have outlined in my study, the intervention does not need to be extreme. For example, a programme which helped parents engage in crafts, songs, indoor games, outdoor activities and telling stories to the child would not have to completely change the parent-child dynamic but would certainly increase the second parental input I defined as *parent-child interactions* in Chapter 4. Indeed, examples of family-based interventions that do incorporate these kinds of parent-child activities currently exist, such as the *Fast Track Program* in the US, *Sure Start* in the UK, or *Families and Schools Together (FAST)* in Canada.[142] Determining the optimal approach to increasing this type of parental engagement in a real-world setting is an important goal supported by the evidence in this thesis.

The empirical findings can also inform other observational research in a range of areas. Given that I have identified the ability of parenting behaviours to predict non-cognitive skills, one example of such observational research would be to investigate if similar activities with non-parental figures can achieve comparable outcomes. Alternatively, the strong persistence of skills (measured using my model) indicates the potential for research which aims to identify other factors that explain why early childhood ability is highly predictive of later outcomes.

The potential for future research extends far beyond the methodological and empirical research projects I have presented above. Regardless of the specific path of study that results from my research, the findings from my thesis add valuable evidence and model-based insights to the existing literature and can help to shape future educational policies and practice.

[142] For more information on the *Fast Track Program*, *Sure Start* and *FAST* see Bierman (2002), H. Roberts (2000) and McDonald (2015) respectively.

References

Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools*, *34*(3), 219-228. doi: 10.1002/(SICI)1520-6807(199707)34:3<219::AID-PITS4>3.0.CO;2-J

Aiyagari, S., Greenwood, J., & Seshadri, A. (2002). Efficient Investment in Children. *Journal of Economic Theory*, *102*(2), 290-321. doi: 10.1006/jeth.2001.2852

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The Impact of the Cattell-Horn-Carroll Theory on Test Development and Interpretation of Cognitive and Academic Abilities. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues.* (p. 185-202). New York, NY, US: The Guilford Press.

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 4, p. 1-181). Elsevier. doi: 10.1016/B978-0-444-53444-6.00001-8

Anderson, C., Kim, J.-S., & Keller, B. (2014). Multilevel Modeling of Categorical Response Variables. In *Handbook of International Large-Scale Assessment* (p. 481-519). New York: Chapman and Hall/CRC. doi: 10.1201/b16061

Areepattamannil, S. (2010). Parenting Practices, Parenting Style, and Children's School Achievement. *Psychological Studies*, *55*(4), 283-289. doi: 10.1007/s12646-010-0043-0

Arnold, D. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Marshall, N. A. (2012). The association between preschool children's social functioning and their emergent academic skills. *Early Childhood Research Quarterly*, *27*(3), 376-386. doi: 10.1016/j.ecresq.2011.12.009

Attanasio, O. (2015). The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement and Policies. *Journal of the European Economic Association*, *13*(6), 949-997. doi: doi:10.1111/jeea.12159

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2015). Estimating the Production Function for Human Capital: Results from a Randomized

Control Trial in Colombia. *National Bureau of Economic Research Working Paper Series, No. 20965*. doi: 10.3386/w20965

Baker, M. (2011). Innis Lecture: Universal early childhood interventions: What is the evidence base? *Canadian Journal of Economics/Revue canadienne d'économique*, *44*(4), 1069-1105. doi: 10.1111/j.1540-5982.2011.01668.x

Baker, M., & Milligan, K. (2016). Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries. *Journal of Human Capital*, *10*(4), 399-441. doi: 10.1086/688899

Bar-On, R. (1997). *The Emotional Quotient Inventory (EQ-i): Technical manual.* Toronto, Canada: Multi-Health Systems Inc.

Bar-On, R. (2004). The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description and summary of psychometric properties. In G. Geher (Ed.), *Measuring emotional intelligence: Common ground and controversy.* (p. 115-145). Hauppauge, NY, US: Nova Science Publishers.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, *18*(Suppl), 13-25.

Bar-On, R. (2010). Emotional Intelligence: An Integral Part of Positive Psychology. *South African Journal of Psychology*, *40*(1), 54-62. doi: 10.1177/008124631004000106

Barendse, M. T., Oort, F. J., & Timmerman, M. E. (2015). Using Exploratory Factor Analysis to Determine the Dimensionality of Discrete Responses. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(1), 87-101. doi: 10.1080/10705511.2014.934850

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child development*, *37*(4), 887-907. doi: 10.2307/1126611

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, *75*(1), 43-88.

Baumrind, D. (1971). Current patterns of parental authority. *Developmental psychology*, *4*(1, pt.2), 1-103. doi: http://dx.doi.org/10.1037/h0030372

Baumrind, D. (1978). Parental disciplinary patterns and social competence in children. *Youth & Society*, *9*(3), 239-267. doi: 10.1177/0044118X7800900302

Baumrind, D. (1991). The Influence of Parenting Style on Adolescent Competence and Substance Use. *The Journal of Early Adolescence*, *11*(1), 56-95. doi: 10.1177/0272431691111004

Baumrind, D. (2005). Patterns of parental authority and adolescent autonomy. *New Directions for Child and Adolescent Development*, *2005*(108), 61-69. doi: 10.1002/cd.128

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*(5), 9-49. doi: http://dx.doi.org/10.1086/258724

Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.*

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, *75*(299), 493-517. doi: 10.2307/2228949

Becker, G. S., & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, *87*(6), 1153-1189. doi: 10.1086/260831

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, *4*(3, Part 2), S1-S39. doi: 10.1086/298118

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, *75*(4), 352-365. doi: 10.1086/259291

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238-246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). EQS structural equations program manual. *Encino*, 83-100. (ISBN 1-885898-03-7 1885898029 978-1885898029)

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606. doi: 10.1037/0033-2909.88.3.588

Bergeman, C. S., & Plomin, R. (1988). Parental mediators of the genetic relationship between home environment and infant mental development. *British Journal of Developmental Psychology*, *6*(1), 11-19. doi: 10.1111/j.2044-835X.1988.tb01076.x

Berger, L. M., Paxson, C., & Waldfogel, J. (2009). Income and Child Development. *Children and youth services review*, *31*(9), 978-989. doi: 10.1016/j.childyouth.2009.04.013

Bevaart, F., Mieloo, C. L., Jansen, W., Raat, H., Donker, M. C. H., Verhulst, F. C., & van Oort, F. V. A. (2012). Ethnic differences in problem perception and perceived need for care for young children with problem behaviour. *Journal of Child Psychology and Psychiatry*, *53*(10), 1063-1071. doi: 10.1111/j.1469-7610.2012.02570.x

Bierman, K. (2002). The Implementation of the Fast Track Program: An Example of a Large-Scale Prevention Science Efficacy Trial. *Journal of Abnormal Child Psychology*, *30*(1), 1-17. doi: 10.1023/A:1014292830216

Biroli, P. (2017). *Health and Skill Formation in Early Childhood* (52 JEL Classification: J2). Centre for Household, Income, Labour and Demographic Economics (CHILD) - CCA. doi: 10.2139/ssrn.2877648\

Bono, E. D., Francesconi, M., Kelly, Y., & Sacker, A. (2016). Early Maternal Time Investment and Early Child Outcomes. *The Economic Journal*, *126*(596), F96-F135. doi: 10.1111/ecoj.12342

Borghans, L., Duckworth, A. L., Heckman, J. J., & ter Weel, B. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, *43*(4), 972-1059. doi: 10.3368/jhr.43.4.972

Bowlby, J. (1969). Attachment and Loss: Volume I: Attachment. In *Attachment and Loss: Volume I: Attachment* (2nd ed., Vol. 1, p. 1-401). London: The Hogarth Press and the Institute of Psycho-Analysis. (ISSN 2472-6982)

Bracken, B. A. (2002). *Bracken School Readiness Assessment.* San Antonio, Texas: The Psychological Corporation.

Bracken, B. A. (2004). *The Psychoeducational Assessment of Preschool Children* (3rd ed.). Mahwah, New Jersey & London: Taylor & Francis, Lawrence Erlbaum Associates P.

Bracken, B. A., & Panter, J. E. (2011). *Using the Bracken Basic Concept Scale and Bracken Concept Development Program in the assessment and remediation of young children's concept development* (Vol. 48). doi: 10.1002/pits.20568

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2011). *Inequality During the Early Years: Child Outcomes and Readiness to Learn in Australia, Canada, United Kingdom, and United States* (SSRN Scholarly Paper No. ID 1965137). Rochester, NY: Social Science Research Network.

Bradbury, B., Waldfogel, J., Washbrook, L., & Corak, M. (2015a). *Too Many Children Left Behind.* New York: Russell Sage Foundation.

Bradbury, B., Waldfogel, J., Washbrook, L., & Corak, M. (2015b). *Too Many Children Left Behind (Technical Appendix).* New York: Russell Sage Foundation.

Bradley, R. H., Caldwell, B. M., & Rock, S. L. (1988). Home environment and school performance: A ten-year follow-up and examination of three models of environmental action. *Child development*, *59(4)*, 852-867. doi: 10.2307/1130253

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic Status and Child Development. *Annual Review of Psychology*, *53*(1), 371-399. doi: 10.1146/annurev.psych.53.100901.135233

Brennan, L. M., Shaw, D. S., Dishion, T. J., & Wilson, M. (2012). Longitudinal Predictors of School-Age Academic Achievement: Unique Contributions of Toddler-Age Aggression, Oppositionality, Inattention, and Hyperactivity. *Journal of Abnormal Child Psychology*, *40*(8), 1289-1300. doi: 10.1007/s10802-012-9639-2

Britton, J., & Vignoles, A. (2017). Education production functions. In *Handbook of Contemporary Education Economics* (p. 246-271). Edward Elgar Publishing.

Bronfenbrenner, U. (1979). *The ecology of human development.* Harvard university press.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In *Handbook of child psychology: Theoretical models of human development, Volume*

1, 5th ed. (p. 993-1028). Hoboken, NJ, US: John Wiley & Sons Inc.

Bronfenbrenner, U., & Morris, P. A. (2006). The Bioecological Model of Human Development. In *Handbook of child psychology: Theoretical models of human development, Vol. 1, 6th ed.* (p. 793-828). Hoboken, NJ, US: John Wiley & Sons Inc.

Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology*, *99*(2), 353-395. doi: 10.1086/230268

Brooks-Gunn, J., & Markman, L. B. (2005). The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness. *The Future of Children*, *15*(1), 139-168.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research, Second Edition* (Vol. Second edition). New York: The Guilford Press.

Browne, M. W. (2001). An Overview of Analytic Rotation in Exploratory Factor Analysis. *Multivariate Behavioral Research*, *36*(1), 111-150. doi: 10.1207/S15327906MBR3601_05

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230-258. doi: 10.1177/0049124192021002005

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-Term Memory, Working Memory, and Executive Functioning in Preschoolers: Longitudinal Predictors of Mathematical Achievement at Age 7 Years. *Developmental Neuropsychology*, *33*(3), 205-228. doi: 10.1080/87565640801982312

Carlo, G., McGinley, M., Hayes, R., Batenhorst, C., & Wilkinson, J. (2007). Parenting Styles or Practices? Parenting, Sympathy, and Prosocial Behaviors Among Adolescents. *The Journal of Genetic Psychology*, *168*(2), 147-176. doi: 10.3200/GNTP.168.2.147-176

Carneiro, P. M., Crawford, C., & Goodman, A. (2007). *The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes* (CEE Discussion Paper). Centre for the Economics of Education, LSE.

Carneiro, P. M., & Heckman, J. J. (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, & B. M. Friedman (Eds.), *Inequality in America: What role for human capital policies?* (p. 77-239). Cambridge, MA: MIT Press.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies.* Cambridge ; New York: Cambridge University Press.

Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. In *Contemporary intellectual assessment: Theories, tests, and issues.* (p. 122-130). New York, NY, US: The Guilford Press.

Carvajal, H. H., Parks, J. P., Bays, K. J., Logan, R. A., Lujano, C. I., Page, G. L., & Weaver, K. A. (1991). Relationships between Scores on Wechsler Preschool and Primary Scale of Intelligence—Revised and Stanford-Binet IV. *Psychological Reports*, *69*(1), 23-26. doi: 10.2466/pr0.1991.69.1.23

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, *38*(592), 10.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1-22. doi: 10.1037/h0046743

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245-276. doi: 10.1207/s15327906mbr0102_10

Caucutt, E. M., Lochner, L., & Park, Y. (2017). Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform so Poorly? *The Scandinavian Journal of Economics*, *119*(1), 102-147. doi: 10.1111/sjoe.12195

Center for Human Resource Research. (2000). *1998 NLSY79 Child and Young Adult Data Users Guide.* The Ohio State University Columbus, OH.

Cerda, C. A., Im, M. H., & Hughes, J. N. (2014). Learning-related skills and academic achievement in academically at-risk first graders. *Journal of Applied Developmental Psychology*, *35*(5), 433-443. doi: 10.1016/j.appdev.2014.08.001

Chao, R. K., & Willms, J. D. (2000). *Family Income, Parenting Practices, and Childhood Vulnerability: A Challenge to the "Culture of Poverty" Thesis* (Tech. Rep. Nos. Policy Brief March 2000, No. 9). Fredericton, NB: Canadian Research Institute for Social Policy, University of New Brunswick.

CLS, C. f. L. S. (2014). *Millenium Cohort Study - Ethical review and Consent.* UCL Institute of Education.

Cobb-Clark, D. A., Nicolás Salamanca, & Zhu, A. (2018). Parenting style as an investment in human development. *Journal of Population Economics*, *32*(4), 1315-1352. doi: 10.1007/s00148-018-0703-2

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6edition ed.). London ; New York: Routledge. doi: 10.4324/9780203029053

Cook, J., & Cook, G. (2013). *The World of Children* (3rd ed.). Pearson Higher Education.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653-665. doi: 10.1016/0191-8869(92)90236-I

Cunha, F., & Heckman, J. J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31-47. doi: 10.1257/aer.97.2.31

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *The Journal of Human Resources*, *43*(4), 738-782. doi: 10.3368/jhr.43.4.738

Cunha, F., & Heckman, J. J. (2009). The Economics and Psychology of Inequality and Human Development. *Journal of the European Economic Association*, *7*(2-3), 320-364. doi: 10.1162/JEEA.2009.7.2-3.320

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Chapter 12 Interpreting the Evidence on Life Cycle Skill Formation. In E. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, p. 697-812). Elsevier. doi: 10.1016/S1574-0692(06)01012-9

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*(3), 883-931. doi: 10.3982/ECTA6551

Dahl, G. B., & Lochner, L. (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, *102*(5), 1927-56. doi: 10.1257/aer.102.5.1927

Darling, N., & Steinberg, L. (1993). Parenting style as context: An integrative model. *Psychological bulletin*, *113*(3), 487. doi: 10.1037/0033-2909.113.3.487

Davies, S., Janus, M., Duku, E., & Gaskin, A. (2015). Using the Early Development Instrument to examine cognitive and non-cognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly*, *35*(2nd quarter 2016), 63-75. doi: 10.1016/j.ecresq.2015.10.002

Del Boca, D., Monfardini, C., & Nicoletti, C. (2017). Parental and child time investments and the cognitive development of adolescents. *Journal of Labor Economics*, *35*(2), 565-608. doi: 10.1086/689479

Demie, F. (2018). English language proficiency and attainment of EAL (English as second language) pupils in England. *Journal of Multilingual and Multicultural Development*, *39*(7), 641-653. doi: 10.1080/01434632.2017.1420658

DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the Personality Inventory for DSM-5. *Personality disorders*, *7*(2), 113-123. doi: 10.1037/per0000170

Dickerson, A., & Popli, G. K. (2016). Persistent poverty and children's cognitive development: Evidence from the UK Millennium Cohort Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(2), 535-558. doi: 10.1111/rssa.12128

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, Late or Never? When Does Parental Education Impact Child Outcomes? *The Economic Journal*, *126*(596), F184-F231. doi: 10.1111/ecoj.12356

Dixit, A. K. (2014). *Microeconomics : A very short introduction / Avinash Dixit.* Oxford : Oxford University Press, 2014.

Doepke, M., & Zilibotti, F. (2017). Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, *85*(5), 1331-1371. doi: 10.3982/ECTA14634

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories:

A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, *47*(2), 309-326. doi: 10.1111/j.2044-8317.1994.tb01039.x

Dooley, M., & Stewart, J. (2007). Family income, parenting styles and child behavioural–emotional outcomes. *Health Economics*, *16*(2), 145-162. doi: 10.1002/hec.1142

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428-1446. doi: 10.1037/0012-1649.43.6.1428

Duncan, G. J., & Magnuson, K. (2013). Investing in Preschool Programs. *Journal of Economic Perspectives*, *27*(2), 109-32. doi: 10.1257/jep.27.2.109

Duncan, G. J., & Murnane, R. J. (2011). *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances: Rising Inequality, Schools, and Children's Life Chances.* Russell Sage Foundation. (https://www.jstor.org/stable/10.7758/9781610447515) doi: 10.7758/9781610447515

Duncombe, M. E., Havighurst, S. S., Holland, K. A., & Frankling, E. J. (2012). Psychometric Evaluation of a Brief Parent- and Teacher-Rated Screen for Children at Risk of Conduct Disorder. *Australian Journal of Educational & Developmental Psychology*, *12*, 1-11.

Dunham, M., Mcintosh, D., & Gridley, B. E. (2002). An Independent Confirmatory Factor Analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment*, *20*(2), 152-163. doi: 10.1177/073428290202000204

Dunn, L. M., & Dunn, L. M. (1981). Manual for the peabody picture vocabulary test-revised. *Circle Pines, MN: American Guidance Service.*

Ehrenberg, R. G. (2017). *Modern labor economics : Theory and public policy* (13th edition. ed.; R. S. Smith, Ed.). London : Routledge. doi: 10.4324/9781315101798

Elliott, C., Smith, P., & McCulloch, K. (1996). *British Ability Scales Second Edition (BAS II): Administration and Scoring Manual.* London: Nelson.

Erikson, E. H. (1950). *Childhood and society.* New York, NY, US: W W Norton & Co.

Erikson, E. H. (1968). *Identity: Youth and crisis.* Oxford, England: Norton & Co.

Ermisch, J. (2008). Origins of Social Immobility and Inequality: Parenting and Early Child Development. *National Institute Economic Review*, *205*(1), 62-71. doi: 10.1177/0027950108096589

Eysenck, H. J. (1991). Dimensions of personality: 16, 5 or 3?—Criteria for a taxonomic paradigm. *Personality and Individual Differences*, *12*(8), 773-790. doi: 10.1016/0191-8869(91)90144-Z

Eysenck, H. J. (1994). The Big Five or giant three: Criteria for a paradigm. In *The* *developing structure of temperament and personality from infancy to adulthood.* (p. 37-51). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory Factor Analysis.* New York, NY: Oxford University Press. doi: 10.1093/acprof:osobl/9780199734177.001.0001

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299. doi: 10.1037/1082-989X.4.3.272

Feinstein, L. (2000). *The relative economic importance of academic, psychological and behavioural attributes developed on childhood* [Monograph]. http://cep.lse.ac.uk.

Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, *70*(277), 73-97. doi: 10.1111/1468-0335.t01-1-00272

Fleishman, E. A. (1967). Development of a behavior taxonomy for describing human tasks: A correlational-experimental approach. *Journal of Applied Psychology*, *51*(1), 1-10. doi: 10.1037/h0024073

Fleishman, E. A., & Bartlett, C. J. (1969). Human abilities. *Annual Review of Psychology*, *20*(1), 349-380. doi: 10.1146/annurev.ps.20.020169.002025

Flouri, E., Midouhas, E., & Joshi, H. (2014). Family Poverty and Trajectories of Children's Emotional and Behavioural Problems: The Moderating Roles of Self-Regulation and Verbal Cognitive Ability. *Journal of Abnormal Child Psychology*, *42*(6), 1043-1056. doi: 10.1007/s10802-013-9848-3

Flouri, E., Midouhas, E., & Ruddy, A. (2016). Socio-economic status and family structure differences in early trajectories of child adjustment: Individual and neighbourhood effects. *Health & Place*, *37*, 8-15. doi: 10.1016/j.healthplace.2015.11.005

Flouri, E., & Sarmadi, Z. (2016). Prosocial behavior and childhood trajectories of internalizing and externalizing problems: The role of neighborhood and school contexts. *Developmental Psychology*, *52*(2), 253-258. doi: 10.1037/dev0000076

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics*, *82*(4), 540-554. doi: 10.1162/003465300559037

Fraenkel, J. R., & Wallen, N. E. (2009). *How to design and evaluate research in education* (7th ed.). McGraw-Hill.

Francesconi, M., & Heckman, J. J. (2016). Child Development and Parental Investment: Introduction. *The Economic Journal*, *126*(596), F1-F27. doi: 10.1111/ecoj.12388

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2016). Are fit indices really fit to estimate the number of factors with categorical variables? Some cautionary findings via Monte Carlo simulation. *Psychological methods*, *21*(1), 93-111. doi: 10.1037/met0000064

Girard, L.-C., Pingault, J.-B., Doyle, O., Falissard, B., & Tremblay, R. E. (2016). Developmental Associations Between Conduct Problems and Expressive Language in Early Childhood: A Population-Based Study. *Journal of Abnormal Child Psychology*, *44*(6), 1033-1043. doi: 10.1007/s10802-015-0094-8

Goldberg, L. R. (1971). *A historical survey of personality scales and inventories.* Oregon Research Institute.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26-34. doi: 10.1037/0003-066X.48.1.26

Goleman, D. (1995). *Emotional intelligence.* New York: New York : Bantam Books.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581-586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(11), 1337-1345. doi: 10.1097/00004583-200111000-00015

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*, *177*(6), 534-539. doi: 10.1192/bjp.177.6.534

Gorard, S. (2001). *Quantitative Methods in Educational Research: The Role of Numbers Made Easy.* A&C Black.

Guo, G., & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography*, *37*(4), 431-447. doi: 10.1353/dem.2000.0005

Gutman, L. M., & Feinstein, L. (2010). Parenting behaviours and children's development from infancy to early childhood: Changes, continuities and contributions. *Early Child Development and Care*, *180*(4), 535-556. doi: 10.1080/03004430802113042

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*(2), 149-161. doi: 10.1007/BF02289162

Hair, J. F. (2018). *Multivariate data analysis / Joseph F. Hair, Jr., William C. Black, Barry J. Babin, Rolph E. Anderson.* (Eighth edition. ed.; W. C. Black, B. J. Babin, & R. E. Anderson, Eds.). Andover : Cengage Learning, 2018.

Hansen, K. (2014). *Millennium Cohort Study A Guide to the Datasets (Eighth Edition)* (Tech. Rep.). University of London. Institute of Education. Centre for Longitudinal Studies.

Harris, M., & Westermann, G. (2014). *A Student's Guide to Developmental Psychology* (1st edition ed.). Taylor & Francis. doi: 10.4324/9781315867212

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD, US: Paul H Brookes Publishing.

Hartas, D. (2011). Children's language and behavioural, social and emotional difficulties and prosocial behaviour during the toddler years and at school entry. *British Journal of Special Education*, *38*(2), 83-91. doi: 10.1111/j.1467-8578.2011.00507.x

Hartas, D. (2012). Inequality and the home learning environment: Predictions about seven-year-olds' language and literacy. *British Educational Research Journal*, *38*(5), 859-879. doi: 10.1080/01411926.2011.588315

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, *33*(4), 1829-1878.

Hayashi, F. (2000). *Econometrics.* Princeton: Princeton University Press. (https://press.princeton.edu/titles/6946.html)

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science (New York, N.Y.)*, *312*(5782), 1900-1902. doi: 10.1126/science.1128898

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, *19*(4), 451-464. doi: 10.1016/j.labeco.2012.05.014

Heckman, J. J., & Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual review of economics*, *6*, 689-733. doi: 10.1146/annurev-economics-080213-040753

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*(3), 411-482. doi: 10.1086/504455

Hedeker, D. (2008). Multilevel Models for Ordinal and Nominal Variables. In J. de Leeuw & E. Meijer (Eds.), *Handbook of Multilevel Analysis* (p. 237-274). Springer New York, NY. doi: 10.1007/978-0-387-73186-5_6

Helmers, C., & Patnam, M. (2011). The formation and evolution of childhood skill acquisition: Evidence from India. *Journal of Development Economics*, *95*(2), 252-266. doi: 10.1016/j.jdeveco.2010.03.001

Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement*, *66*(3), 393-416. doi: 10.1177/0013164405282485

Hernández-Alava, M., & Popli, G. (2017). Children's Development and Parental Input: Evidence From the UK Millennium Cohort Study. *Demography*, *54*(2), 485-511. doi: 10.1007/s13524-017-0554-6

Hill, V. (2005). Through the Past Darkly: A Review of the British Ability Scales Second Edition. *Child and Adolescent Mental Health*, *10*(2), 87-98. doi: 10.1111/j.1475-3588.2004.00123.x

Hofstee, W. K., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*(1), 146-163. doi: 10.1037/0022-3514.63.1.146

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185. doi: 10.1007/BF02289447

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological review*, *75*(3), 242. doi: 10.1037/h0025662

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. *Woodcock-Johnson technical manual*, 197-232.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of educational psychology*, *57*(5), 253. doi: 10.1037/h0023816

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In *Contemporary intellectual assessment: Theories, tests, and issues.* (p. 53-91). New York, NY, US: The Guilford Press.

Horn, J. L., & Stankov, L. (1982). Auditory and visual factors of intelligence. *Intelligence*, *6*(2), 165-185. doi: 10.1016/0160-2896(82)90012-5

Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014). Predicting first-grade mathematics achievement: The contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in Psychology*, *5*, 272. doi: 10.3389/fpsyg.2014.00272

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, *24*(6), 417. doi: 10.1037/h0071325

Howard-Jones, P., Washbrook, E., & Meadows, S. (2012). The timing of educational investment: A neuroscientific perspective. *Neuroscience & Education*, *2*, S18-S29. doi: 10.1016/j.dcn.2011.11.002

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications, Third Edition.* Taylor & Francis. doi: 10.4324/9781315650982

Hsin, A., & Felfe, C. (2014). When Does Time Matter? Maternal Employment, Children's Time With Parents, and Child Development. *Demography*, *51*(5), 1867-1894. doi: 10.1007/s13524-014-0334-5

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, *3*(4), 424. doi: 10.1037/1082-989X.3.4.424

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. doi: 10.1080/10705519909540118

In'nami, Y., & Rie Koizumi. (2013). Structural Equation Modeling in Educational Research: A Primer. In (p. 23-51). Leiden, The Netherlands: Brill | Sense. doi: 10.1007/9789462093324_003

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter Estimates: Some

Support for the N:q Hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 128-141. doi: 10.1207/S15328007SEM1001_6

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. *Handbook of personality: Theory and research.*, 66-100.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research, 2nd ed.*, 102-138. (Guilford Publications, 1999. https://pages.uoregon.edu/sanjay/pubs/bigfive.pdf)

Jöreskog, K. G., & Sörbom, D. (1982). Recent Developments in Structural Equation Modeling. *Journal of Marketing Research*, *19*(4), 404-416. doi: 10.2307/3151714

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, *20*(1), 141-151. doi: 10.1177/001316446002000116

Kalil, A. (2015). Inequality Begins at Home: The Role of Parenting in the Diverging Destinies of Rich and Poor Children. In P. R. Amato, A. Booth, S. M. McHale, & J. Van Hook (Eds.), *Families in an Era of Increasing Inequality: Diverging Destinies* (p. 63-82). Cham: Springer International Publishing. doi: 10.1007/978-3-319-08308-7_5

Kaufman, J. C., Kaufman, S. B., & Plucker, J. A. (2013). Contemporary Theories of Intelligence. In Daniel Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (p. 811-822). doi: 10.1093/oxfordhb/9780195376746.013.0051

Keenan, T., Evans, S., & Crowley, K. (2016). *An Introduction to Child Development* (3rd ed.). London: SAGE Publications Ltd.

Kelloway, K. E. (2015). *Using Mplus for Structural Equation Modeling: A Researcher's Guide* (Second Edition ed.). Thousand Oaks, California: SAGE Publications. doi: 10.4135/9781483381664

Kelly, Y., Sacker, A., Del Bono, E., Francesconi, M., & Marmot, M. (2011). What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study. *Archives of Disease in Childhood*, *96*(9), 832-837. doi: 10.1136/adc.2010.195917

Kiernan, K. E., & Huerta, M. C. (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood1. *The British Journal of Sociology*, *59*(4), 783-806. doi: 10.1111/j.1468-4446.2008.00219.x

Kiernan, K. E., & Mensah, F. K. (2011). Poverty, family resources and children's early educational attainment: The mediating role of parenting. *British Educational Research Journal*, *37*(2), 317-336. doi: 10.1080/01411921003596911

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition* (4th ed.). New York, United States: Guilford Publications.

Koo, A., & Chan, T. W. (2010). Parenting Style and Youth Outcomes in the UK*. *European Sociological Review*, *27*(3), 385-399. doi: 10.1093/esr/jcq013

Ladd, G. W., & Price, J. M. (1987). Predicting Children's Social and School Adjustment Following the Transition from Preschool to Kindergarten. *Child Development*, *58*(5), 1168-1189. doi: 10.2307/1130613

Lagerspetz, K. M. J., Björkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? gender differences in aggressiveness in 11- to 12-year-old children. *Aggressive Behavior*, *14*(6), 403-414. doi: 10.1002/1098-2337(1988)14:6<403::AID-AB2480140602>3.0.CO;2-D

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*(3), 441-451. doi: 10.1016/j.paid.2006.08.001

Lamborn, S. D., Mounts, N. S., Steinberg, L., & Dornbusch, S. M. (1991). Patterns of Competence and Adjustment among Adolescents from Authoritative, Authoritarian, Indulgent, and Neglectful Families. *Child Development*, *62*(5), 1049-1065. doi: 10.1111/j.1467-8624.1991.tb01588.x

Landy, S., & Tam, K. K. (1996). Yes, parenting does make a difference to the development of children in Canada. *Growing Up in Canada: National Longitudinal Survey of Children and Youth*, 117-126.

Landy, S., & Tam, K. K. (1998). Understanding the contribution of multiple risk factors on child development at various ages. In *Working Paper W-98-22E.* Human Resources Development Canada, Applied Research Branch.

Lareau, A. (2011). *Unequal childhoods class, race, and family life: With an update a decade later.* (2nd edition ed.). Berkeley, Calif.: University of California Press.

Laughlin, T. (1995). The School Readiness Composite of the Bracken Basic Concept Scale as an Intellectual Screening Instrument. *Journal of Psychoeducational Assessment*, *13*(3), 294-302. doi: 10.1177/073428299501300308

Lee, S. M., Daniels, M. H., & Kissinger, D. B. (2006). Parental Influences on Adolescent Adjustment: Parenting Styles Versus Parenting Practices. *The Family Journal*, *14*(3), 253-259. doi: 10.1177/1066480706287654

Lefebvre, P., Merrigan, P., & Verstraete, M. (2008). The Effects of School Quality and Family Functioning on Youth Math Scores: A Canadian Longitudinal Analysis. *SSRN Electronic Journal*, 32. doi: 10.2139/ssrn.1266686

Leibowitz, A. (1974). Home investments in children. *Journal of Political Economy*, *82*(2, Part 2), S111-S131. doi: 10.1086/260295

Lohr, S. (2009). *Sampling: Design and Analysis* (2nd ed.). Cengage Learning.

Love, J. M., Kisker, E. E., Ross, C., Raikes, H., Constantine, J., Boller, K., ... Vogel, C. (2005). The effectiveness of early head start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental psychology*, *41*(6), 885-901. doi: 10.1037/0012-1649.41.6.88

Lubke, G. H., & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural equation modeling*, *11*(4), 514-534. doi: 10.1207/s15328007sem1104_2

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, *4*(1), 84. doi: 10.1037/1082-989X.4.1.84

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interaction. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.) (Eds.), *Handbook of child psychology: Vol. 4: Socialization, personality and social development* (4th ed ed., p. 1-101). New York: Wiley.

Mankiw, N. G. (2015). *Principles of microeconomics / N. Gregoy Mankiw.* (Seventh edition. ed.). Stamford, Connecticut : Cengage Learning, [2015].

Mas-Colell, A. (1995). *Microeconomic theory.* (J. R. Green & M. D. Whinston, Eds.). New York: Oxford University Press, New York.

Matte, T. D., Bresnahan, M., Begg, M. D., & Susser, E. (2001). Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: Cohort study. *BMJ: British Medical Journal*, *323*(7308), 310-314. doi: 10.1136/bmj.323.7308.310

Mavroveli, S., Petrides, K. V., Shove, C., & Whitehead, A. (2008). Investigation of the construct of trait emotional intelligence in children. *European Child & Adolescent Psychiatry*, *17*(8), 516-526. doi: 10.1007/s00787-008-0696-6

Mayer, J. D. (2002). MSCEIT: Mayer-salovey-caruso emotional intelligence test. *Toronto, Canada: Multi-Health Systems.*

Mayer, J. D. (2008). Personal Intelligence. *Imagination, Cognition and Personality*, *27*(3), 209-232. doi: 10.2190/IC.27.3.b

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267-298. doi: 10.1016/S0160-2896(99)00016-1

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The Ability Model of Emotional Intelligence: Principles and Updates. *Emotion Review*, *8*(4), 290-300. doi: 10.1177/1754073916639667

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2012). Does Personal Intelligence Exist? Evidence From a New Ability-Based Measure. *Journal of Personality Assessment*, *94*(2), 124-140. doi: 10.1080/00223891.2011.646108

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human Abilities: Emotional Intelligence. *Annual Review of Psychology*, *59*(1), 507-536. doi: 10.1146/annurev.psych.59.103006.093646

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? *Emotional development and emotional intelligence: Educational implications.*, 3-34.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion (Washington, D.C.)*, *3*(1), 97-105. doi: 10.1037/1528-3542.3.1.97

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947-959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly*, *15*(3), 307-329. doi: 10.1016/S0885-2006(00)00069-7

McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, *60*(2), 175-215. doi: 10.1111/j.1467-6494.1992.tb00970.x

McCulloch, A., & Joshi, H. E. (2002). Child development and family resources: Evidence from the second generation of the 1958 British birth cohort. *Journal of Population Economics*, *15*(2), 283-304. doi: 10.1007/s001480100067

McDonald, L. (2015). Families and schools together (FAST). In J. J. Ponzetti Jr. (Ed.), *Evidence-based Parenting Education: A Global Perspective* (p. 244-264). New York, NY: Routledge. doi: 10.4324/9781315766676

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In *Contemporary intellectual assessment: Theories, tests, and issues.* (p. 151-119). New York, NY, US: The Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1-10. doi: 10.1016/j.intell.2008.08.004

McLoyd, V. (1998). Socioeconomic disadvantage and child development. *The American psychologist*, *53*(2), 185-204. doi: 10.1037//0003-066x.53.2.185

McWayne, C. M., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool Competency in Context: An Investigation of the Unique Contribution of Child Competencies to Early Academic Success. *Developmental Psychology*, *40*(4), 633-645. doi: 10.1037/0012-1649.40.4.633

Meadows, S. (2017). *Understanding Child Development: Psychological Perspectives and*

Applications (2nd ed.). Routledge, an imprint of Taylor & Francis.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2000). *The Mental Health of Children and Adolescents in Great Britain: (622732007-001).* American Psychological Association. doi: 10.1037/e622732007-001

Michaud, S. (2001). The National Longitudinal Survey of Children and Youth - Overview and Changes After Three Cycles. *Canadian Studies in Population*, *28*(2), 391-405. doi: 10.25336/P6QK58

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, *66*(4), 281-302.

Mistry, R. S., Biesanz, J. C., Chien, N., Howes, C., & Benner, A. D. (2008). Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. *Early Childhood Research Quarterly*, *23*(2), 193-212. doi: 10.1016/j.ecresq.2008.01.002

Morongiello, B. A. (1997). *Tapping School Readiness in the NLSCY: Measurement Issues and Solutions.* (Tech. Rep. No. T-98-1E). Ottawa: Applied Research Branch, Strategic Policy: Human Resources and Development Canada. (Departmental catalogue number T-98-1E)

Murnane, R. J., Willett, J. B., & Levy, F. (1995). The Growing Importance of Cognitive Skills in Wage Determination. *The Review of Economics and Statistics*, *77*(2), 251-266. doi: 10.3386/w5076

Muthén, L., & Muthén, B. (2017). *Mplus User's Guide* (Eighth ed.). Los Angeles: Muthén & Muthén.

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2008). Best Practices in Exploratory Factor Analysis. In *Best Practices in Quantitative Methods* (p. 86-99). Thousand Oaks, California: SAGE Publications, Inc. doi: 10.4135/9781412995627

Pagani, L. S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: A French Canadian replication and extension. *Developmental Psychology*, *46*(5), 984-994. doi: 10.1037/a0018881

Panter, J. E. (2000). Validity of the Bracken Basic Concept Scale-Revised for Predicting Performance on the Metropolitan Readiness Test-Sixth Edition. *Journal of Psychoeducational Assessment*, *18*(2), 104-110. doi: 10.1177/073428290001800201

Panter, J. E., & Bracken, B. A. (2009). Validity of the Bracken School Readiness Assessment for predicting first grade readiness. *Psychology in the Schools*, *46*(5), 397-409. doi: 10.1002/pits.20385

Paro, K. M. L., & Pianta, R. C. (2000). Predicting Children's Competence in the Early School Years: A Meta-Analytic Review. *Review of Educational Research*, *70*(4), 443-484. doi: 10.3102/00346543070004443

Paulson, S. E. (1994). Relations of Parenting Style and Parental Involvement with

Ninth-Grade Students' Achievement. *The Journal of Early Adolescence*, *14*(2), 250-267. doi: 10.1177/027243169401400208

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559-572. doi: 10.1080/14786440109462720

Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality*, *15*(6), 425-448. doi: 10.1002/per.416

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioural validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality*, *17*(1), 39-57. doi: 10.1002/per.466

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology*, *98*(2), 273-289. doi: 10.1348/000712606X120618

Petriwskyj, A., Thorpe, K., & Tayler, C. (2005). Trends in construction of transition to school in three western regions, 1990–2004. *International Journal of Early Years Education*, *13*(1), 55-69. doi: 10.1080/09669760500048360

Phillips, D. A., & Shonkoff, J. P. (2000). *From neurons to neighborhoods: The science of early childhood development.* National Academies Press. (https://www.ncbi.nlm.nih.gov/books/NBK225557/)

Piaget, J. (1952a). *The origins of intelligence in children.* (2nd ed.; M. Cook, Ed.). New York, NY, US: W W Norton & Co. doi: 10.1037/11494-000

Piaget, J. (1952b). *Play, dreams and imitation in childhood.* New York, NY, US: W W Norton & Co.

Piaget, J. (1954). *The construction of reality in the child.* (M. Cook, Ed.). New York, NY, US: Basic Books. doi: 10.1037/11168-000

Piérard, E., Buckley, N., & Chowhan, J. (2004). Bootstrapping Made Easy: A Stata ADO File. *Information and Technical Bulletin*, *1*(1), 20-36.

Plewis, I. (2014). *The Millennium Cohort Study: Technical Report on Sampling (Fourth Edition).* UCL. Institute of Education. Centre for Longitudinal Studies.

Poulou, M. S. (2014). How are trait emotional intelligence and social skills related to emotional and behavioural difficulties in adolescents? *Educational Psychology*, *34*(3), 354-366. doi: 10.1080/01443410.2013.785062

Poulou, M. S. (2015). Teacher-Student Relationships, Social and Emotional Skills, and Emotional and Behavioural Difficulties. *International Journal of Educational Psychology*, *4*(1), 25.

Poulou, M. S. (2017). An examination of the relationship among teachers' perceptions of social-emotional learning, teaching efficacy, teacher-student interactions, and students' behavioral difficulties. *International Journal of School & Educational Psychology*, *5*(2), 126-136. doi: 10.1080/21683603.2016.1203851

Rasbash, J., Leckie, G., Pillinger, R., & Jenkins, J. (2010). Children's educational progress: Partitioning family, school and area effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *173*(3), 657-682. doi: 10.1111/j.1467-985X.2010.00642.x

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354-373. doi: 10.1037/a0029315

Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An Ecological Perspective on the Transition to Kindergarten: A Theoretical Framework to Guide Empirical Research. *Journal of Applied Developmental Psychology*, *21*(5), 491-511. doi: 10.1016/S0193-3973(00)00051-4

Roberts, H. (2000). What is Sure Start? *Archives of disease in childhood*, *82*(6), 435-437. doi: 10.1136/adc.82.6.435

Roberts, R. D., Schulze, R., & MacCann, C. (2008). The measurement of emotional intelligence: A decade of progress? In *The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing.* (Vol. 2, p. 461-482). Thousand Oaks, CA, US: Sage Publications, Inc. doi: 10.4135/9781849200479.n22

Ronalds, G. A., De Stavola, B. L., & Leon, D. A. (2005). The cognitive cost of being a twin: Evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *Bmj*, *331*(7528), 1306. doi: 10.1136/bmj.38633.594387.3A

Rutter, M., & Redshaw, J. (1991). Annotation: Growing up as a Twin: Twin-Singleton Differences in Psychological Development. *Journal of Child Psychology and Psychiatry*, *32*(6), 885-895. doi: 10.1111/j.1469-7610.1991.tb01916.x

Sabol, T. J., & Pianta, R. C. (2012). Patterns of School Readiness Forecast Achievement and Socioemotional Development at the End of Elementary School. *Child Development*, *83*(1), 282-299. doi: 10.1111/j.1467-8624.2011.01678.x

Salovey, P., & Mayer, J. D. (1990). Emotional Intelligence. *Imagination, Cognition and Personality*, *9*(3), 185-211. doi: 10.2190/DUGG-P24E-52WK-6CDG

Santrock, J. W. (2017). *Life-span development* (16th Edition ed.). New York, NY: McGraw-Hill Education.

Schmitt, T. A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment*, *29*(4), 304-321. doi: 10.1177/0734282911406653

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In *Contemporary intellectual assessment: Theories, tests, and issues, 3rd ed.* (p. 99-

144). New York, NY, US: The Guilford Press.

Schoon, I. (2010). Childhood cognitive ability and adult academic attainment: Evidence from three British cohort studies. *Longitudinal and Life Course Studies*, *1*(3), 241-158. doi: 10.14301/llcs.v1i3.93

Schoon, I., Jones, E., Cheng, H., & Maughan, B. (2012). Family hardship, family instability, and cognitive development. *Journal of Epidemiology and Community Health*, *66*(8), 716-722. doi: 10.1136/jech.2010.121228

Shaffer, D. R. (2014). *Developmental psychology: Childhood and adolescence.* (9th edition, International edition ed.; K. Kipp, Ed.). Belmont, CA: Cengage Learning.

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563-575. doi: 10.1007/BF02296196

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy.* Piscataway, NJ, US: Transaction Publishers.

Sparrow, S. S., & Davis, S. M. (2000). Recent Advances in the Assessment of Intelligence and Cognition. *Journal of Child Psychology and Psychiatry*, *41*(1), 117-131. doi: 10.1111/1469-7610.00552

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201-292. doi: 10.2307/1412107

Spera, C. (2005). A Review of the Relationship Among Parenting Practices, Parenting Styles, and Adolescent School Achievement. *Educational Psychology Review*, *17*(2), 125-146. doi: 10.1007/s10648-005-3950-1

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2018). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*, *90*(4), 1308-1318. doi: 10.1111/cdev.13072

Spira, E. G., & Fischel, J. E. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: A review. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*(7), 755-773. doi: 10.1111/j.1469-7610.2005.01466.x

Statistics Canada. (1996). *Growing up in Canada. National Longitudinal Survey of Children and Youth. Survey Overview.* Minister of Industry. (Location: Ottawa)

Statistics Canada. (2004). *Microdata User Guide: National Longitudinal Survey of Children and Youth Cycle 8.* Special Surveys Division: HRSDC.

Statistics Canada. (2010). *Mitigation of Risk to Respondents of Statistics Canada's Surveys.* Research Data Centres Program.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, *25*(2), 173-180. doi: 10.1207/s15327906mbr2502_4

Steiger, J. H. (2016). Notes on the Steiger–Lind (1980) Handout. *Structural Equation*

Modeling: A Multidisciplinary Journal, *23*(6), 777-781. doi: 10.1080/10705511.2016.1217487

Steinberg, L., Lamborn, S. D., Dornbusch, S. M., & Darling, N. (1992). Impact of Parenting Practices on Adolescent Achievement: Authoritative Parenting, School Involvement, and Encouragement to Succeed. *Child Development*, *63*(5), 1266-1281. doi: 10.1111/j.1467-8624.1992.tb01694.x

Stock, J. H., & Watson, M. W. (2002). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, *97*(460), 1167-1179. doi: 10.1198/016214502388618960

Stone, L. L., Janssens, J. M. A. M., Vermulst, A. A., Van Der Maten, M., Engels, R. C. M. E., & Otten, R. (2015). The Strengths and Difficulties Questionnaire: Psychometric properties of the parent and teacher version in children aged 4–7. *BMC Psychology*, *3*(1), 12. doi: 10.1186/s40359-015-0061-8

Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric Properties of the Parent and Teacher Versions of the Strengths and Difficulties Questionnaire for 4- to 12-Year-Olds: A Review. *Clinical Child and Family Psychology Review*, *13*(3), 254-274. doi: 10.1007/s10567-010-0071-2

Strand, S., Malmberg, L., & Hall, J. (2015). English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database.

Strayhorn, J. M., & Weidman, C. S. (1988). A Parent Practices Scale and Its Relation to Parent and Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(5), 613-618. doi: 10.1097/00004583-198809000-00016

Thurstone, L. (1947). *Multiple factor analysis.* University of Chicago Press: Chicago.

Thurstone, L. (1959). *The Measurement of Values.* Chicago: University of Chicago Press.

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement*. *The Economic Journal*, *113*(485), F3-F33. doi: 10.1111/1468-0297.00097

Todd, P. E., & Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, *1*(1), 91-136. doi: 10.1086/526401

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1-10. doi: 10.1007/BF02291170

Vogler, P., Crivello, G., & Woodhead, M. (2008). *Early Childhood Transitions Research: A Review of Concepts, Theory, and Practice. Working Papers in Early Childhood Development, No. 48.* Bernard van Leer Foundation.

Vygotsky, L. S. (1962). *Thought and language.* (E. Hanfmann & G. Vakar, Eds.).

Cambridge, MA, US: MIT Press. doi: 10.1037/11193-000

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). London, England: Harvard University Press.

Waldfogel, J. (2007). Parental Work Arrangements and Child Development. *Canadian Public Policy*. doi: 10.3138/cpp.33.2.251

Washbrook, E., Propper, C., & Sayal, K. (2013). Pre-school hyperactivity/attention problems and educational outcomes in adolescence: Prospective longitudinal study. *The British Journal of Psychiatry*, *203*(4), 265-271. doi: 10.1192/bjp.bp.112.123562

Weir, K., & Duveen, G. (1981). Further Development and Validation of the Prosocial Behaviour Questionnaire for use by Teachers. *Journal of Child Psychology and Psychiatry*, *22*(4), 357-374. doi: 10.1111/j.1469-7610.1981.tb00561.x

Welsh, M., Parke, R. D., Widaman, K., & O'Neil, R. (2001). Linkages Between Children's Social and Academic Competence: A Longitudinal Analysis. *Journal of School Psychology*, *39*(6), 463-482. doi: 10.1016/S0022-4405(01)00084-X

Whiteside-Mansell, L., Bradley, R. H., Tresch Owen, M., Randolph, S. M., & Cauce, A. M. (2003). Parenting and children's behavior at 36 months: Equivalence between African American and European American mother - child dyads. *Parenting: Science and Practice*, *3*(3), 197-234. doi: 10.1207/S15327922PAR0303_02

Winter, D. G., & Barenbaum, N. B. (1999). History of modern personality theory and research. *Handbook of personality: Theory and research, 2nd ed.*, 3-27. (ISBN 9781572304833)

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, *73*(6), 913-934. doi: 10.1177/0013164413495237

Yang, Y., & Xia, Y. (2015). On the number of factors to retain in exploratory factor analysis for ordered categorical data. *Behavior Research Methods*, *47*(3), 756-772. doi: 10.3758/s13428-014-0499-2

Yeung, W. J., Linver, M. R., & Brooks–Gunn, J. (2002). How Money Matters for Young Children's Development: Parental Investment and Family Processes. *Child Development*, *73*(6), 1861-1879. doi: 10.1111/1467-8624.t01-1-00511

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes /* (Unpublished doctoral dissertation). University of California, Los Angeles, Los Angeles, CA.

Zwirs, B. W. C., Burger, H., Buitelaar, J. K., & Schulpen, T. W. J. (2006). Ethnic differences in parental detection of externalizing disorders. *European Child & Adolescent Psychiatry*, *15*(7), 418-426. doi: 10.1007/s00787-006-0550-7

APPENDIX A

Additional Details about MCS Data and Analysis

A.1 Summary of MCS Survey Elements

Provided below are Tables 19-23 as they are included in the Millenium Cohort Study: A Guide to the Datasets . Each table provides detail on the measures included in one sweep of the MCS.

Table 19: MCS1 – Summary of MCS1 Survey Elements.

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

Table 20: MCS2 – Summary of MCS2 Survey Elements

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

Table 21: MCS3 – Summary of MCS3 Survey Elements

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

Table 22: MCS4 – Summary of MCS4 Survey Elements.

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

Table 23: MCS5 – Summary of FIFTH Survey Elements.

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

Table 23: MCS5 – Summary of FIFTH Survey Elements (continued from previous page)

Table describing MCS survey elements removed for copyright reasons.
Copyright holder is Institute of Education.

For original table, refer to Hansen (2014)

A.2 MCS Sampling Strategies

To adequately capture various subgroups, the target sample for the MCS was designed so that it oversampled certain areas of the country where the census data indicated the presence of certain ethnic minorities and low-income populations (Plewis, 2014). This was achieved through a modified random sample, which made use of the statistical techniques of stratification and clustering.

Stratification is a sampling technique that divides a population into distinct groups using measurable characteristics and then takes a random sample from each of these groups. For example, a population could be divided based on gender and then a random sample be taken within each gender. By using the stratification process, researchers can make sure that an adequately sized sample is collected from each group. This differs from a simple random sample, which by design has a non-zero probability of creating a sample which does not include a given subgroup.

The MCS was stratified on two levels. First, the population of the UK was divided by country, and then within each country the CLS used existing information about the composition of electoral wards in order to assign classifications to each electoral ward. These gave the strata used by the MCS. The exact definition varied by country. In England, the electoral wards were assigned one of three classifications: ethnic minority wards were those where ethnic minorities formed at least 30 percent of the population in the 1991 Census; disadvantaged wards were those in the bottom quartile as defined by the Child Poverty Index; and advantaged wards captured all wards that met neither of the other criteria. In Wales, Scotland and Northern Ireland, only the advantaged and disadvantaged strata were used.

In a simple stratified sample, researchers would randomly sample within each of the defined strata. The MCS instead made use of a technique known as cluster sampling. Cluster sampling involves selecting a number of subgroups and then collecting data only from these groups in order to represent the whole population (Cohen et al., 2007). This is often done when it would be logistically or financially unfeasible to use random sampling. In the MCS, the cluster sampling took the form of randomly selecting electoral wards within each of the defined strata and then collecting data on the entire population of interest within these wards. More specifically, the MCS randomly selected wards and then collected data on all births that took place in each of these wards during the relevant time period. This choice of sample design reduced the research costs and provided the added ability to compare various electoral wards over time.

A.3 Full Code for Analysis

Presented below is the MPlus input file used to obtain the results presented in Section 4.4.

```
TITLE: "PhD Analysis - U.K. Empirical Application"

DATA: FILE IS "/Users/Ashton/Desktop/Millenium Cohort/mpluscleaned.dat";

VARIABLE: NAMES ARE SPTN00 EOVWT2 PTTYPE2
BSRT3 BASnvT3 BASnvT5 BASpsT5 BASpcT5 NFERT7 BASwrT7 BASpcT7 BASvsT11
BSRP3 BASnvA3 BASnvA5 BASpsA5 BASpcA5 NFERA7 BASwrA7 BASpcA7 BASvsA11
SDQPr3 SDQSo3 SDQCo3  SDQEm3 SDQHy3
SDQPr5 SDQSo5 SDQCo5  SDQEm5 SDQHy5
SDQPr7 SDQSo7 SDQCo7  SDQEm7 SDQHy7
SDQPr11 SDQSo11 SDQCo11  SDQEm11 SDQHy11
lib3 reads3  music3  art3  alph3  count3
lib5 reads5  music5 pgame5  ind5 park5 stors5 art5
count5  writ5  read5
lib7 reads7 pgame7  ind7  park7  stors7  music7 art7
count7 writ7  read7
white langEO langOth male
birthwgt mageb meduc_1 meduc_2 meduc_3 meduc_4 meduc_5 meduc_6 meduc_7
inc1Q_2 inc1Q_3 inc1Q_4 inc1Q_5  inc3Q_2 inc3Q_3 inc3Q_4 inc3Q_5
inc5Q_2 inc5Q_3 inc5Q_4 inc5Q_5  inc7Q_2 inc7Q_3 inc7Q_4 inc7Q_5
bmonth sib1 sib3 sib5 sib7
singp1 singp3 singp5 singp7 agem1 agem3 agem5 agem7 agem11 ;

VARIABLE: USEVARIABLES ARE SPTN00 EOVWT2 PTTYPE2
BSRP3 BASnvA3 BASnvA5 BASpsA5 BASpcA5 NFERA7 BASwrA7 BASpcA7 BASvsA11
SDQPr3 SDQSo3 SDQCo3  SDQEm3 SDQHy3
SDQPr5 SDQSo5 SDQCo5  SDQEm5 SDQHy5
SDQPr7 SDQSo7 SDQCo7  SDQEm7 SDQHy7
SDQPr11 SDQSo11 SDQCo11  SDQEm11 SDQHy11
lib3 reads3  music3  art3  alph3  count3
lib5 reads5  music5 pgame5  ind5 park5 stors5 art5
count5  writ5  read5
lib7 reads7 pgame7  ind7  park7  stors7  music7 art7
count7 writ7  read7
white langEO langOth male
birthwgt mageb meduc_1 meduc_2 meduc_3 meduc_4 meduc_5 meduc_6 meduc_7
inc1Q_2 inc1Q_3 inc1Q_4 inc1Q_5  inc3Q_2 inc3Q_3 inc3Q_4 inc3Q_5
inc5Q_2 inc5Q_3 inc5Q_4 inc5Q_5  inc7Q_2 inc7Q_3 inc7Q_4 inc7Q_5
bmonth sib3 sib5 sib7
singp3 singp5 singp7  agem3 agem5 agem7 agem11 ;
```

```
CATEGORICAL ARE lib3 reads3  music3  art3  alph3  count3
lib5 reads5 music5 pgame5  ind5 park5 stors5 art5
count5  writ5  read5
lib7 reads7 pgame7  ind7  park7  stors7  music7 art7
count7 writ7  read7  ;

MISSING = . ;

WEIGHT IS EOVWT2;
STRATIFICATION IS PTTYPE2;
CLUSTER IS SPTN00;

ANALYSIS: TYPE = COMPLEX;
ITERATIONS = 10000;

MODEL:
lcog3 BY BSRP3 BASnvA3 ;
lcog5 BY BASnvA5 BASpsA5 BASpcA5 ;
lcog7 BY NFERA7 BASwrA7 BASpcA7 ;

lnoncog3 BY SDQPr3@-1 SDQSo3 SDQCo3 SDQEm3 SDQHy3;
lnoncog5 BY SDQPr5@-1 SDQSo5 SDQCo5 SDQEm5 SDQHy5;
lnoncog7 BY SDQPr7@-1 SDQSo7 SDQCo7 SDQEm7 SDQHy7;
lnoncog11 BY SDQPr11@-1 SDQSo11 SDQCo11 SDQEm11 SDQHy11;

linv3r BY reads3 lib3;
linv3m BY music3 art3 alph3 count3;

linv5r BY reads5 lib5  ;
linv5m BY music5 art5 pgame5 ind5 park5 stors5 ;
linv5a BY count5 writ5 read5 ;

linv7r BY reads7 lib7 ;
linv7m BY music7 art7 pgame7 ind7 park7 stors7 ;
linv7a BY count7 writ7 read7;

lcog5 ON lcog3 lnoncog3 linv3m linv3r ;
lcog7 ON lcog5 lnoncog5 linv5m linv5r linv5a ;
BASvsA11 ON lcog7 lnoncog7 linv7m linv7r linv7a ;

lnoncog5 ON lcog3 lnoncog3 linv3m linv3r ;
lnoncog7 ON lcog5 lnoncog5 linv5m linv5r linv5a ;
lnoncog11 ON lcog7 lnoncog7 linv7m linv7r linv7a ;

lcog3 ON birthwgt mageb
meduc_1 meduc_2 meduc_3 meduc_4 meduc_5 meduc_6 meduc_7 bmonth ;
lnoncog3 ON birthwgt mageb
```

```
meduc_1 meduc_2 meduc_3 meduc_4 meduc_5 meduc_6 meduc_7 bmonth ;

lcog3 ON   inc1Q_2 inc1Q_3 inc1Q_4 inc1Q_5 bmonth;
lcog5 ON   inc3Q_2 inc3Q_3 inc3Q_4 inc3Q_5 bmonth;
lcog7 ON   inc5Q_2 inc5Q_3 inc5Q_4 inc5Q_5 bmonth ;
BASvsA11 ON  inc7Q_2 inc7Q_3 inc7Q_4 inc7Q_5 bmonth  ;

lnoncog3 ON  inc1Q_2 inc1Q_3 inc1Q_4 inc1Q_5 bmonth ;
lnoncog5 ON  inc3Q_2 inc3Q_3 inc3Q_4 inc3Q_5 bmonth ;
lnoncog7 ON  inc5Q_2 inc5Q_3 inc5Q_4 inc5Q_5 bmonth ;
lnoncog11 ON inc7Q_2 inc7Q_3 inc7Q_4 inc7Q_5 bmonth ;

BSRP3   ON agem3 white langEO langOth male ;
BASnvA3 ON agem3 white langEO langOth male ;
BASpsA5 ON agem5 white langEO langOth male ;
BASpcA5 ON agem5 white langEO langOth male ;
BASnvA5 ON agem5 white langEO langOth male ;
BASpcA7 ON agem7 white langEO langOth male ;
BASwrA7 ON agem7 white langEO langOth male ;
NFERA7  ON agem7 white langEO langOth male ;
BASvsA11 ON agem11 white langEO langOth male ;

SDQPr3 ON white male ;
SDQSo3 ON white male ;
SDQCo3 ON white male ;
SDQEm3 ON white male ;
SDQHy3 ON white male ;
SDQPr5 ON white male ;
SDQSo5 ON white male ;
SDQCo5 ON white male ;
SDQEm5 ON white male ;
SDQHy5 ON white male ;
SDQPr7 ON white male ;
SDQSo7 ON white male ;
SDQCo7 ON white male ;
SDQEm7 ON white male ;
SDQHy7 ON white male ;
SDQPr11 ON white male ;
SDQSo11 ON white male ;
SDQCo11 ON white male ;
SDQEm11 ON white male ;
SDQHy11 ON white male ;

linv3m ON sib3 singp3;
linv3r ON sib3 singp3;

linv5m ON sib5 singp5;
```

```
linv5r ON sib5 singp5;
linv5a ON sib5 singp5;

linv7m ON sib7 singp7;
linv7r ON sib7 singp7;
linv7a ON sib7 singp7;

lcog3 WITH lcog5@0 lcog7@0 BASvsA11@0
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv3m@0 linv3r@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

lcog5 WITH lcog7@0 BASvsA11@0
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv3m@0 linv3r@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

lcog7 WITH BASvsA11@0
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv3m@0 linv3r@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

BASvsA11 WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv3m@0 linv3r@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

linv3m WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

linv3r WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv5m@0 linv5r@0 linv5a@0
linv7m@0 linv7r@0 linv7a@0;

linv5m WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv7m@0 linv7r@0 linv7a@0;

linv5r WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
```

```
linv7m@0 linv7r@0 linv7a@0;

linv5a WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0
linv7m@0 linv7r@0 linv7a@0;

linv7m WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0 ;

linv7r WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0 ;

linv7a WITH
lnoncog3@0 lnoncog5@0 lnoncog7@0 lnoncog11@0 ;

lnoncog3 WITH
lnoncog5@0 lnoncog7@0 lnoncog11@0 ;

lnoncog5 WITH
lnoncog7@0 lnoncog11@0 ;

lnoncog7 WITH
lnoncog11@0 ;

OUTPUT: TECH4 STDY STDYX;
```

APPENDIX B

Additional Information about the NLSCY

B.1 Summary of NLSCY Survey Elements

On the following page is a summary of the general survey elements included in each cycle of the NLSCY originally presented by Michaud (2001).The exact contents of the NLSCY data varies between cycles. Therefore, this summary should only be used to provide a general sense of the types of variables included in the NLSCY. Specific details about each cycle of the NLSCY are available in the documentation provided by Statistics Canada.

Links to the questionnaires, user guides and survey methodologies can be found at: **https://crdcn.org/datasets/nlscy-national-longitudinal-survey-children-and-youth**.

Figure outlining the survey elements of the NLSCY removed for copyright reasons.

Copyright holder is Statistics Canada.

For original figure, refer to Michaud (2001)

Fig. B.1 Summary of NLSCY Measures – figured obtained from p. 401 of Michaud (2001)

B.2 Application for Access to Canadian Data

Statistics Canada has chosen to limit access to certain confidential data to secure locations that are known as Research Data Centres (RDCs). These data centres are housed within Canadian universities had have strict regulations in regards to access and use of data. In order to access the data as a visiting student, a local researcher had to support the project and an application was submited to the Social Sciences and Humanities Research Council (SSHRC). Below is a copy of this application.

Social Sciences and Humanities Research Council of Canada | Statistics Canada

Conseil de recherches en sciences humaines du Canada | Statistique Canada

Internal use
943390

Research Data Centre Application

Identification

This page will be made available to the RDC Access Granting commitee.

Funding Opportunity

Access to Statistics Canada Research Data Centres

Project title

A dynamic model of skill development: empirically modelling the role of parental investment in the development of cognitive and non-cognitive skills in longitudinal samples of primary school children

Keywords

cognitive skills, non-cognitive skills, skill formation, parental investment

Do you have funding for this project? (●) Yes ( ) No

Sources of funding

Other sources of funding

Canada-UK Foundation, Cambridge Commonwealth Trust, University of Cambridge Faculty of Education

Applicant family name	Applicant given name	Initials
Brown	Ashton	

Research Data Centre Location	Access period
University of Calgary - Prairie RDC	Start date 2017/9 End date 2018/6

Reference Letter

If you are applying as a student, indicate the name of the author of your reference letter

Author family name	Author given name	Initials

Surveys

NLSCY - National Longitudinal Survey of Children and Youth / NLSCY all cycles (C1: 1994-1995 to C8: 2008-2009)

Family name, Given name
Brown, Ashton

Team Members

List the names of your research team members. Indicate the location of the Research Data Centre each team member will use.

Family name		Given name	Initials
Org. code	Full organization name		
Department/Division name			
Research Data Centre location			
Family name		Given name	Initials
Org. code	Full organization name		
Department/Division name			
Research Data Centre location			
Family name		Given name	Initials
Org. code	Full organization name		
Department/Division name			
Research Data Centre location			
Family name		Given name	Initials
Org. code	Full organization name		
Department/Division name			
Research Data Centre location			
Family name		Given name	Initials
Org. code	Full organization name		
Department/Division name			
Research Data Centre location			

RDC Project Proposal
July 1, 2017
Ashton Brown
University of Cambridge

1 Title

A dynamic model of skill development: empirically modelling the role of parental investment in the development of cognitive and non-cognitive skills in primary school children

2 Rationale and Objectives of the Study

This proposal seeks access to the microdata files for Cycles 1-8 of the National Longitudinal Survey of Children and Youth (NLSCY) in order to write the fourth and final chapter of my PhD dissertation. My PhD is a four-part project that uses existing longitudinal data in order to examine the trajectories of cognitive and non-cognitive skill development in primary school children. The Canadian data would provide a comparative sample for the work I've already done on the United Kingdom using the longitudinal data contained in the Millennium Cohort Study (MCS). Access to the NLSCY will allow me to measure the trajectories of skill development in Canadian children, an area which is previously unexplored in the literature.

There is substantial interest in understanding the developmental trajectories of cognitive and non-cognitive skills in primary school children and the role that parental input plays in the development of these skills. This research is motivated by a desire to identify the behaviours, skills and traits in early childhood that will lead to socio-economic success in adulthood. If it is possible to identify the parenting behaviours that promote the development of skills it allows policy makers create interventional programs to identify children who are at risk and propose interventions.

Research Questions
This project will allow me to address the following questions:

1. How do the trajectories of cognitive and non-cognitive skills in Canadian children compare to the trajectories of these skills in children in the United Kingdom
2. What role does parental input play in skill development and are early childhood skills a determinant of the level of parental investment in later periods of childhood?

3 Methods

In order to examine the trajectories that these skills take, I have proposed a modification to the currently accepted theoretical framework for empirically measuring skill development. This modification allows the model to capture how parental investment changes over time in response to the child's development.The proposed model is the combination of two existing theoretical frameworks: 'The Ecological and Dynamic Model of Transition' presented by

Rimm-Kaufman and Pianta (2000), and the economic 'Model of Skill Formation' originally presented by Cunha and Heckman (2007). The former describes skill development using the transition to formal education, while the latter models it as an economic production function, similar to other economic investments. I argue that neither model fully captures the nature of childhood skill development, but by drawing on aspects of both models, we will be better able to understand the mechanisms that underlie this process.

I have already shown the potential of my proposed methodology using data from the United Kingdom, and found that it is able to provide more detailed analysis of skill development in primary school children. The findings were consistent across the United Kingdom and and hope to use Canadian data to provide an international comparison.

3.1 Empirically Modelling Skill Formation

The main analytical framework of my analysis is a two part process. A full theoretical explanation is beyond the scope of this proposal, but a detailed exposition of the mathematical assumptions that justify this approach is available upon request. Below I provide a brief summary of the analysis.

Creating the Latent Variables:

The first step of the analysis the use of maximum likelihood estimation to construct the latent variables of cognitive skills, non-cognitive skills and parental investment behaviours. For each of these I have identified observable measures, and I estimate the factor loadings using maximum likelihood equations. These are obtained using maximum likelihood estimation in Stata with the following equations:

For cognitive ability, θ^C, two cognitive test scores, $CogScore$, are used to predict a latent cognitive score, for each Cycle of the NLSCY. This equation takes the form:

$$\hat{\theta}_t^C = \beta_0^C + \beta_1^C CogScore1_t + \beta_2^C CogScore2_t + \eta_t^C \tag{1}$$

A similar equation with the parent reported behaviour measures is used to estimate latent non-cognitive skill, θ^N, for the child in each cycle as:

$$\hat{\theta}_t^N = \beta_0^N + \beta_1^N B1_t + \beta_2^N B2_t + \beta_3^N B3_t + +\beta_4^N B4_t + \beta_5^N B5_t + \eta_t^N \tag{2}$$

Similarly, the investment measures are based on activities the parent reports engaging with the child, for example reading to the child, participating in arts and crafts, singing songs and attending the library. This can be used to provide latent estimates of parental input, which is denoted I:

$$\hat{I}_t = \beta_0^I + \beta_1^I Read_t + \beta_2^I Art_t + \beta_3^I Music_t + \beta_4^I Library_t + \eta_t^I \tag{3}$$

Then, using the estimated values of β obtained with the above regressions, as well as the observations from each individual, I am able to estimate latent scores for each of the relevant measures at the four time points. These estimated latent scores will provide the basis for the remainder of my analysis.

Measuring the Trajectory of Skills Over Time:

Once I have created these latent scores, they are then used to conduct repeated regressions using an extension of the linear specification of Cunha and Heckman (2008). To extend the model, I have included additional variables to measure the trajectory of parental behaviours, denoted as I. The system of equations is as follows:

$$\begin{pmatrix} \theta^N_{t+1} \\ \theta^C_{t+1} \\ I_{t+1} \end{pmatrix} = \begin{pmatrix} \gamma^N_1 & \gamma^N_2 & \gamma^N_3 \\ \gamma^C_1 & \gamma^C_2 & \gamma^C_3 \\ \gamma^I_1 & \gamma^I_2 & \gamma^I_3 \end{pmatrix} \begin{pmatrix} \theta^N_t \\ \theta^C_t \\ I_t \end{pmatrix} + \begin{pmatrix} \eta^N \\ \eta^C_t \\ \eta^I_t \end{pmatrix} \tag{4}$$

which can be reduced to a linear law of motion for non-cognitive, cognitive skills and investment as shown below:

$$\theta^N_{t+1} = \gamma^N_0 + \gamma^N_1 \theta^N_t + \gamma^N_2 \theta^C_t + \gamma^N_3 I_t + \eta^N_t \tag{5}$$

$$\theta^C_{t+1} = \gamma^C_0 + \gamma^C_1 \theta^N_t + \gamma^C_2 \theta^C_t + \gamma^C_3 I_t + \eta^C_t \tag{6}$$

$$I_{t+1} = \gamma^I_0 + \gamma^I_1 \theta^N_t + \gamma^I_2 \theta^C_t + \gamma^I_3 I_t + \eta^I_t \tag{7}$$

Each of these equations represents a regression that will be conducted using Stata. The coefficients γ indicate the role that the various inputs play in future skill growth. The coefficients also measure the role that a child's skill in one period plays in predicting parental behaviour in the next period.

4 Data Requirements

The analysis for this project requires access to Cycles 1-8 of the National Longitudinal Survey of Children and Youth. The project will take children who were ages 0-3 at the time of Cycle 1 and track them over time. The longitudinal NLSCY data is needed because the proposed research requires following respondents from early-childhood through to late-adolescence. The project I am proposing requires access to the master files as the public use data lacks the child identifiers needed to link the respondents across cycles. The expected sample size is quite large, as the relevant sub-sample of children is over 6,000 respondents during Cycle 1. This analysis will not use small levels of geography and the final analysis will be on a national scale.

This project will look at two sets of variables, those examining family characteristics that might influence skill development, and those measuring skills and behaviours over time.

For the family characteristics, I will look primarily at the data reported by the PMK. The list is extensive, but the two series which are especially relevant are parental education, *EDPQ*, and parental income *INHQ*. I have not listed the demographic variables, but the analysis will include variables for child-age, gender, birthweight, mother's age at birth, etc.

The skills and behaviours over time will require extensive measures so it is not possible to list all the variables. However, I intend to focus on certain types of questions. For non-cognitive skills I will use the PMK's responses to the behavioural symptoms checklist *ECQ*. Cognitive skills will be assessed using the Peabody Picture Vocabulary Test (PPVT)

PPCS01, Who Am I Test *WICdS01*, Number Knowledge Test *KNCdS01* and Canadian Achievement Tests *MACS02 and RECS02*. For parenting behaviours the relevant questions will vary by Cycle, but I intend to look at the parenting style questions which are contained in the *RLCQ, PRCB, PRCbS* series of variables.

5 Expected Output

The main product of this study will be a chapter of my PhD Dissertation to fulfill the requirements of the University of Cambridge's Doctoral program. I also hope to submit a condensed version of this chapter as an article for peer-reviewed research journals.

6 Proposed Period of Research

September 2017 - December 2017

7 References

This list of references is selective and only lists the references directly cited in this proposal. It does not completely list the works which were referenced in designing this research project.

American Psychiatric Association. 1994. *Diagnostic and Statistical Manual of Mental Disorders, 4th ed.* Arlington, VA: American Psychiatric Press.

Cunha, F., and Heckman, J. J. (2008). *Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation.* The Journal of Human Resources, 43(4), 738?782.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2005). *Interpreting the Evidence on Life Cycle Skill Formation* (Working Paper No. 11331). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w11331

Goodman, R. (1997). *The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry*, 38(5), 581?586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Rimm-Kaufman, S. E., and Pianta, R. C. (2000). *An Ecological Perspective on the Transition to Kindergarten: A Theoretical Framework to Guide Empirical Research.* Journal of Applied Developmental Psychology, 21(5), 491?511. https://doi.org/10.1016/S0193-3973(00)00051-4

UNIVERSITY OF
CALGARY

Prairie Regional Research Data Centre

University of Calgary
MacKimmie Library Tower 116
Telephone: (403) 210-5483
Fax: (403) 282-9298
Email: wanner@ucalgary.ca

July 11, 2017

Mika Oehling
RDC Program Officer
SSHRC
350 Albert Street
P.O. Box 1610
Ottawa, ON
K1P 6G4

Dear Ms. Oehling:

I am writing on behalf of Ashton Brown, a Canadian PhD student at Cambridge University, who will be submitting a proposal to access confidential Statistics Canada data in the Prairie Regional RDC at the University of Calgary. The project title is: *A dynamic model of skill development: empirically modelling the role of parental investment in the development of cognitive and non-cognitive skills in longitudinal samples of British and Canadian children*. Having reviewed her proposal, I find it suitable for submission to the RDC program and fully support Ms. Brown's application for access to our RDC.

Ashton Brown is prepared to analyze data in the RDC as she has completed both undergraduate and graduate courses in econometrics, and data analysis. She also has research experience working with Statistics Canada data as a research assistant on another RDC project as well as her MA thesis at the University of Ottawa. In addition, she has experience with both the Stata and SPSS software packages that will be utilized in the research project.

Several levels of support will be available to Ms. Brown should she need any additional resources while working in the PRRDC. As the local RDC academic director, I will be available to come to the Centre to assist with issues that arise and to preview output before vetting. Her PhD supervisor at the University of Cambridge will be available for questions and support via email and telephone. Ashton has already spoken with her supervisor about methodological considerations and has worked with her in conducting a similar analysis using British data with Stata. As well, our experienced analysts are always available on site to assist researchers with many aspects of data analysis and interpretation.

Sincerely,

Richard A. Wanner, PhD
Academic Director

Director
Microdata Access Division
Statistics Canada
R.H. Coats Building, Floor 9 R
100 Tunney's Pasture Driveway,
Ottawa ON K1A 0T6

1 August 2017

Dear Director of Microdata Access Divison

I would like to inform you that Ashton Brown of the University of Cambridge has submitted an application for access to Statistics Canada data. I have known Ashton Brown for 3 years and know that she is a researcher in good standing at the University of Cambridge. I confirm that she is conducting research in their area of expertise where there is a legitimate requirement to access confidential Canadian data. I am confident she will abide by the procedures in place to protect the confidentiality of respondents of Statistics Canada data.

Sincerely,

Anna Vignoles
Professor of Education (1938)
Director of Research

184 Hills Road, Cambridge,
CB2 8PQ, UK
+ 44 (0)1223 767626

For Office Use Only - Insert all applicable contract numbers

Contract #
Contract #
Contract #
Contract #
Contract #

B.3 NLSCY Sampling Strategies

Instead of constructing an entirely new sampling structure, the NLSCY drew on the existing Statistics Canada Labour Force Survey (LFS). The LFS is a monthly household survey designed to provide estimates of employment, and unemployment in Canada. The survey is designed to provide a representative sample of the Canadian population, and follows selected households for a period of six months. The LFS is a rolling survey and each month respondents who have been involved for six months of the survey are dropped and a new group is selected to replace them.

Respondents to the LFS are selected using a stratified, multistage probability method. More specifically, each province is treated as a separate population from which a stratified and clustered sample is drawn. Stratification means that the provinces are divided into regions based on urban or rural areas, and each of these is divided into clusters such as neighbourhood or city block. This strategy guarantees that the LFS surveys individuals from each cluster and provides a representative picture of the Canadian labour market.

The LFS includes demographic questions which identify the age of members living in each household. Responses from the LFS were used to identify households with children and these were then included in the NLSCY sample. Because the NLSCY sample was identified through the LFS, discussions of the sample design are generally dependent on the sample design of the LFS.

B.4 Survey Weights

The use of clustering and stratification mean that the NLSCY is not a simple random sampling of the Canadian population and contains unequal representation representation of children in the sample. This complex sample design means that the unweighted distribution of characteristics observed in the NLSCY sample may not reflect the distribution of these characteristics in the reference population. The NLSCY is also subject to item non-response and attrition. For these reasons, unweighted estimates of population parameters using the NLSCY data may not provide results that reflect the reference population.

In order to allow researchers to make meaningful inferences about the characteristics of the reference population, Statistics Canada provides final survey weights. These weights not only account for the sampling design, and attrition, but are also adjusted so that they sum to match known population totals. This final adjustment is known as post-stratification. Using these weights in statistical analysis provides statistical estimates that reflect the distribution of characteristics within the reference population.

Although using the final design weights provided by the NLSCY in regression analysis will yield statistical estimates that reflect the distribution of the reference population, the reported standard errors for these estimates are not a reliable estimate of the sampling variance and often overestimate the statistical significance of the estimate.

B.5 Bootstrap Weights

The complex sample design, non-response adjustments and post-stratification of the NLSCY make it impossible to calculate sampling variance using traditional statistical methods. Instead variance estimates are calculated with the use of bootstrap weights.

Bootstrap weights can either be calculated using statistical software or entered directly from the data. In addition to the sampling weights discussed above, Statistics Canada also provides a set of 1,000 bootstrap weights for each respondent. These bootstrap weights account for the specific complex sample design of the NLSCY.

Re-estimating the model of interest separately using each of these bootstrap weights provide 1,000 estimates for each parameter. The variance of these 1,000 estimates provides a reliable estimate of the sampling variance for the estimate obtained using the original sample weight. This estimate of sampling variance will in turn allow for the calculation of test statistics which correctly establish statistical significance.

This bootstrap weighting process can also be used to calculate reliable variance estimates for summary statistics. Unless otherwise noted, all analysis presented in this dissertation is weighted using the longitudinal funnel weights and standard errors which are bootstrapped.

The bswreg command in Stata streamlines this process and uses the bootstrap weights provided in the NLSCY to calculate reliable variance estimates. See Piérard, Buckley, and Chowhan (2004) for a detailed explanation of the use of the bswreg command as well as the reliability of its estimates. Piérard et al. (2004) provide the .ado file required for the use of this command in Appendix I of their paper.

In MPlus, bootstrap weights are built into the original analysis design and can be applied using the bootstrap command. See Muthén and Muthén (2017) for specific details about the use of bootstrap weights in MPlus.

B.6 Full Code for Analysis

Presented below is the MPlus input file used to obtain the results presented in Section 5.4.

```
TITLE: "PhD Thesis - NLSCY Data"

DATA: FILE IS "P:\Brown_5244\Code\MPlus\analysissample.dat" ;

VARIABLE: NAMES ARE adjwgt6 funwt6 bsw1-bsw1000
iPrais2 iFivMin2 iLaugh2 iSpeci2 iSport2
iAnnyd2 iPrPrai2 iPrDis2 iAnger2 iMood2 iProbs2 iRepDis2
iEnf2 iPunis2 iGetawa2 iAvPuni2 iIgPuni2
iPrais3 iFivMin3 iLaugh3 iSpeci3 iSport3
iAnnyd3 iPrPrai3 iPrDis3 iAnger3 iMood3 iProbs3 iRepDis3
iEnf3 iPunis3 iGetawa3 iAvPuni3 iIgPuni3
iPrais4 iFivMin4 iLaugh4 iSpeci4 iSport4
iAnnyd4 iPrPrai4 iPrDis4 iAnger4 iMood4 iProbs4 iRepDis4
iEnf4 iPunis4 iGetawa4 iAvPuni4 iIgPuni4
iPrais5 iFivMin5 iLaugh5 iSpeci5 iSport5
iAnnyd5 iPrPrai5 iPrDis5 iAnger5 iMood5 iProbs5 iRepDis5
iEnf5 iPunis5 iGetawa5 iAvPuni5 iIgPuni5
hscore2 edascr2 pascore2 sascore2
hscore3 edascr3 pascore3 iascore3 poscore3
hscore4 edascr4 pascore4 iascore4
hscore5 edascr5 pascore5 iascore5 poscore5
hscore6 edascr6 pascore6 iascore6 poscore6
cog2 cog3 cog4 cog5 cog6
zcog2 zcog3 zcog4 zcog5 zcog6
xcog2 xcog3 xcog4 xcog5 xcog6
incC1_q1 incC1_q2 incC1_q3 incC1_q4 incC1_q5
incC2_q1 incC2_q2 incC2_q3 incC2_q4 incC2_q5
incC3_q1 incC3_q2 incC3_q3 incC3_q4 incC3_q5
incC4_q1 incC4_q2 incC4_q3 incC4_q4 incC4_q5
incC5_q1 incC5_q2 incC5_q3 incC5_q4 incC5_q5
female white NonNtSpk lowbwgt mageb sib1 sib2 sib3 sib4 sib5
singmom1 singmom2 singmom3 singmom4 singmom5
peduc6_1 peduc6_2 peduc6_3 peduc6_4 peduc6_5 peduc6_6
agem1 agem2 agem3 pagem3 agem4 agem5 agem6 monthb ppvtC4 ;

VARIABLE: USEVARIABLES ARE funwt6
iPrais3 iFivMin3 iLaugh3 iSpeci3 iSport3
iAnnyd3 iPrPrai3 iPrDis3 iAnger3 iMood3 iProbs3 iRepDis3
iEnf3 iPunis3 iGetawa3 iAvPuni3 iIgPuni3
iPrais4 iFivMin4 iLaugh4 iSpeci4 iSport4
iAnnyd4 iPrPrai4 iPrDis4 iAnger4 iMood4 iProbs4 iRepDis4
iEnf4 iPunis4 iGetawa4 iAvPuni4 iIgPuni4
```

```
iPrais5 iFivMin5 iLaugh5 iSpeci5 iSport5
iAnnyd5 iPrPrai5 iPrDis5 iAnger5 iMood5 iProbs5 iRepDis5
iEnf5 iPunis5 iGetawa5 iAvPuni5 iIgPuni5
hscore3 edascr3 pascore3 iascore3 poscore3
hscore4 edascr4 pascore4 iascore4
hscore5 edascr5 pascore5 iascore5 poscore5
hscore6 edascr6 pascore6 iascore6 poscore6
xcog3 xcog4 xcog5 xcog6
incC2_q2 incC2_q3 incC2_q4 incC2_q5
incC3_q2 incC3_q3 incC3_q4 incC3_q5
incC4_q2 incC4_q3 incC4_q4 incC4_q5
incC5_q2 incC5_q3 incC5_q4 incC5_q5
peduc6_2 peduc6_3 peduc6_4 peduc6_5 peduc6_6
female white NonNtSpk lowbwgt mageb sib3 sib4 sib5
singmom3 singmom4  singmom5
agem3 pagem3 agem4 agem5 agem6 ppvtC4    ;

CATEGORICAL ARE
iPrais3 iFivMin3 iLaugh3 iSpeci3 iSport3
iAnnyd3 iPrPrai3 iPrDis3 iAnger3 iMood3 iProbs3 iRepDis3
iEnf3 iPunis3 iGetawa3 iAvPuni3 iIgPuni3
iPrais4 iFivMin4 iLaugh4 iSpeci4 iSport4
iAnnyd4 iPrPrai4 iPrDis4 iAnger4 iMood4 iProbs4 iRepDis4
iEnf4 iPunis4 iGetawa4 iAvPuni4 iIgPuni4
iPrais5 iFivMin5 iLaugh5 iSpeci5 iSport5
iAnnyd5 iPrPrai5 iPrDis5 iAnger5 iMood5 iProbs5 iRepDis5
iEnf5 iPunis5 iGetawa5 iAvPuni5 iIgPuni5 ;

MISSING = . ;
WEIGHT IS funwt6;
REPWEIGHTS ARE bsw1-bsw1000 ;

ANALYSIS:
TYPE=COMPLEX;
REPSE = BOOTSTRAP;
ITERATIONS = 10000;

MODEL:
linv3a BY iPrais3 iFivMin3 iLaugh3 iSpeci3 iSport3 ;
linv3b BY iAnnyd3 iPrPrai3 iPrDis3 iAnger3 iMood3 iProbs3 iRepDis3 ;
linv3c BY iEnf3 iPunis3 iGetawa3 iAvPuni3 iIgPuni3 ;

linv4a BY iPrais4 iFivMin4 iLaugh4 iSpeci4 iSport4 ;
linv4b BY iAnnyd4 iPrPrai4 iPrDis4 iAnger4 iMood4 iProbs4 iRepDis4 ;
linv4c BY iEnf4 iPunis4 iGetawa4 iAvPuni4 iIgPuni4 ;

linv5a BY iPrais5 iFivMin5 iLaugh5 iSpeci5 iSport5 ;
```

```
linv5b BY iAnnyd5 iPrPrai5 iPrDis5 iAnger5 iMood5 iProbs5 iRepDis5 ;
linv5c BY iEnf5 iPunis5 iGetawa5 iAvPuni5 iIgPuni5 ;

lnoncog3 BY hscore3@-1 edascr3 pascore3 iascore3 poscore3  ;
lnoncog4 BY hscore4@-1 edascr4 pascore4 iascore4  ;
lnoncog5 BY hscore5@-1 edascr5 pascore5 iascore5 poscore5 ;
lnoncog6 BY hscore6@-1 edascr6 pascore6 iascore6 poscore6 ;

xcog4 ON xcog3 lnoncog3 linv3a linv3b linv3c ppvtC4 ;
xcog5 ON xcog4 lnoncog4 linv4a linv4b linv4c ;
xcog6 ON xcog5 lnoncog5 linv5a linv5b linv5c ;

lnoncog4 ON xcog3 lnoncog3 linv3a linv3b linv3c ;
lnoncog5 ON xcog4 lnoncog4 linv4a linv4b linv4c ;
lnoncog6 ON xcog5 lnoncog5 linv5a linv5b linv5c ;

xcog3 ON incC2_q2 incC2_q3 incC2_q4 incC2_q5
female white NonNtSpk pagem3 lowbwgt mageb
peduc6_2 peduc6_3 peduc6_4 peduc6_5 peduc6_6  ;

xcog4 ON incC3_q2 incC3_q3 incC3_q4 incC3_q5
female white NonNtSpk agem4  ;

xcog5 ON incC4_q2 incC4_q3 incC4_q4 incC4_q5
female white NonNtSpk agem5 ;

xcog6 ON incC5_q2 incC5_q3 incC5_q4 incC5_q5
female white NonNtSpk agem6 ;

lnoncog3 ON incC2_q2 incC2_q3 incC2_q4 incC2_q5 lowbwgt mageb
peduc6_2 peduc6_3 peduc6_4 peduc6_5 peduc6_6  ;
lnoncog4 ON incC3_q2 incC3_q3 incC3_q4 incC3_q5 ;
lnoncog5 ON incC4_q2 incC4_q3 incC4_q4 incC4_q5  ;
lnoncog6 ON incC5_q2 incC5_q3 incC5_q4 incC5_q5  ;

linv3a on sib3 singmom3 ;
linv3b on sib3 singmom3 ;
linv3c on sib3 singmom3 ;

linv4a on sib4 singmom4 ;
linv4b on sib4 singmom4 ;
linv4c on sib4 singmom4 ;

linv5a on sib5 singmom5 ;
linv5b on sib5 singmom5 ;
linv5c on sib5 singmom5 ;
```

```
hscore3 ON white female agem3 ;
edascr3 ON white female agem3 ;
pascore3 ON white female agem3 ;
iascore3 ON white female agem3 ;
poscore3 ON white female agem3 ;
hscore4 ON white female agem4 ;
edascr4 ON white female agem4 ;
pascore4 ON white female agem4 ;
iascore4 ON white female agem4 ;
hscore5 ON white female agem5 ;
edascr5 ON white female agem5 ;
pascore5 ON white female agem5 ;
iascore5 ON white female agem5 ;
poscore5 ON white female agem5 ;
hscore6 ON white female agem6 ;
edascr6 ON white female agem6 ;
pascore6 ON white female agem6 ;
iascore6 ON white female agem6 ;
poscore6 ON white female agem6 ;

xcog3 WITH xcog4@0 xcog5@0 xcog6@0
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3a@0 linv3b@0 linv3c@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

xcog4 WITH xcog5@0 xcog6@0
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3a@0 linv3b@0 linv3c@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

xcog5 WITH xcog6@0
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3a@0 linv3b@0 linv3c@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

xcog6 WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3a@0 linv3b@0 linv3c@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv3a WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3b@0 linv3c@0
```

```
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv3b WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv3c@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv3c WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv4a@0 linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv4a WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv4b@0 linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv4b WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv4c@0
linv5a@0 linv5b@0 linv5c@0 ;

linv4c WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv5a@0 linv5b@0 linv5c@0 ;

linv5a WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv5b@0 linv5c@0 ;

linv5b WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0
linv5c@0 ;

linv5c WITH
lnoncog3@0 lnoncog4@0 lnoncog5@0 lnoncog6@0 ;

lnoncog3 WITH lnoncog4@0 lnoncog5@0 lnoncog6@0 ;

lnoncog4 WITH lnoncog5@0 lnoncog6@0 ;

lnoncog5 WITH lnoncog6@0 ;

OUTPUT: TECH4 STDY STDYX;
```

B.7 Results for 5 Cycle Model

Table B.1 Cognitive Measures: Covariate Parameter Estimates

	Cognitive Ability (θ^C_{t+1})				
	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Female	0.184 ***	-0.106 ***	-0.113 ***	-0.011	0.003
	(0.010)	(0.013)	(0.011)	(0.012)	(0.012)
White	0.138 ***	0.024	0.046 **	-0.138 ***	0.023
	(0.020)	(0.026)	(0.020)	(0.021)	(0.021)
No English/French Spoken In Home	0.000	-0.235 ***	0.019	0.025	-0.024
	(0.015)	(0.023)	(0.022)	(0.018)	(0.018)
Child Age (in months)	0.662 ***	0.27 **	-0.024	0.495 ***	0.001
	(0.029)	(0.136)	(0.047)	(0.081)	(0.100)

Note: All models are estimated using provided survey weights and bootstrap weights.
*p<0.1 , **p<0.05, ***p<0.01
Standard errors in parentheses.

Table B.2 Parental Investment Measurement Model: Parameter Estimates

Measure	Cycle 2 I_2^1	Cycle 2 I_2^2	Cycle 2 I_2^3	Cycle 3 I_3^1	Cycle 3 I_3^2	Cycle 3 I_3^3	Cycle 4 I_4^1	Cycle 4 I_4^2	Cycle 4 I_4^3	Cycle 5 I_5^1	Cycle 5 I_5^2	Cycle 5 I_5^3
Positive Interaction												
Praise Child	1			1			1			1		
	-			-			-			-		
Five Minutes Focused Attention	1.059			1.109			1.144			1.091		
	(0.047)			(0.027)			(0.033)			(0.023)		
Laugh Together with Child	1.148			1.041			1.131			1.175		
	(0.038)			(0.025)			(0.032)			(0.025)		
Do Something Special with Child	0.747			0.808			0.902			0.918		
	(0.039)			(0.028)			(0.032)			(0.026)		
Play Sports with Child	0.967			0.851			0.928			0.839		
	(0.043)			(0.030)			(0.035)			(0.022)		
Ineffective Parenting												
Get Annoyed with Child		1			1			1			1	
		-			-			-			-	
Proportion of talk: Praise		0.781			0.996			0.924			0.886	
		(0.042)			(0.038)			(0.030)			(0.031)	
Proportion of talk: disapproval		1.135			1.196			1.052			0.974	
		(0.053)			(0.038)			(0.031)			(0.032)	
Angry while punishing child		1.175			0.972			0.946			1.113	
		(0.044)			(0.035)			(0.031)			(0.028)	
Punishment Depends on Mood		1.233			0.879			0.723			0.903	
		(0.063)			(0.038)			(0.033)			(0.027)	
Problems Managing Child		1.405			1.317			1.147			1.164	
		(0.068)			(0.038)			(0.038)			(0.033)	
Need for Repeated Discipline		1.277			1.133			1.103			1.180	
		(0.072)			(0.032)			(0.032)			(0.026)	
Consistent Parenting												
Make sure child follows commands			1			1			1			1
			-			-			-			-
Punish Child if breaks rules			0.924			1.043			1.268			1.078
			(0.055)			(0.035)			(0.041)			(0.052)
Child gets away with breaking rules			1.508			1.355			1.481			1.529
			(0.104)			(0.047)			(0.064)			(0.053)
Child Able to Avoid Punishment			1.317			1.148			1.519			1.436
			(0.089)			(0.043)			(0.078)			(0.052)
Child Ignores Punishment			1.398			1.206			1.494			1.422
			(0.090)			(0.053)			(0.077)			(0.046)

Notes: p<0.01 for all variables
Standard errors shown in parentheses.

B.7 Results for 5 Cycle Model

Table B.3 Measurement Model: Non-Cognitive Skills - Parameter Estimates

PANEL A:	Latent Variables				
	Cycle 2 θ_2^{NC}	Cycle 3 θ_3^{NC}	Cycle 4 θ_4^{NC}	Cycle 5 θ_5^{NC}	Cycle 6 θ_5^{NC}
Hyperactivity-Inattention	-1	-1	-1	-1	-1
	-	-	-	-	-
Emotional Disorder-Anxiety	-0.254	-0.464	-0.400	-0.521	-0.655
	(0.019)	(0.021)	(0.027)	(0.024)	(0.026)
Physical Aggression	-1.177	-0.663	-0.525	-0.699	-0.584
	(0.050)	(0.030)	(0.024)	(0.032)	(0.024)
Separation Anxiety	-0.333				
	(0.026)				
Indirect Aggression		-0.307	-0.32	-0.418	-0.393
		(0.022)	(0.019)	(0.023)	(0.022)
Property Offences		-0.544		-0.492	-0.367
		(0.025)		(0.020)	(0.014)

PANEL B:	Observed Covariates														
	Cycle 2			Cycle 3			Cycle 4			Cycle 5			Cycle 6		
	White	Female	Age in Months	White	Female	Age in Months	White	Female	Age in Months	White	Female	Age in Months	White	Female	Age in Months
Hyperactivity-Inattention	0.839 ***	0.247 ***	-0.04	-0.091	-0.847 ***	-0.47 ***	-0.232	-0.706 ***	-0.093 ***	-0.424 *	-1.191 ***	-0.2 ***	-0.202	-1.12 ***	0.115 ***
	(0.175)	(0.071)	(0.026)	(0.215)	(0.092)	(0.094)	(0.192)	(0.080)	(0.021)	(0.231)	(0.079)	(0.045)	(0.229)	(0.071)	(0.036)
Emotional Disorder-Anxiety	0.036	-0.032	-0.031 ***	0.487 ***	-0.159 ***	-0.212 ***	0.325 **	0.12 **	0.121 ***	0.598 ***	0.042	-0.126 ***	0.241	-0.151 **	0.096 ***
	(0.090)	(0.034)	(0.011)	(0.129)	(0.054)	(0.046)	(0.148)	(0.057)	(0.014)	(0.158)	(0.058)	(0.038)	(0.165)	(0.059)	(0.026)
Physical Aggression	0.338 *	-0.038	-0.12 ***	0.589 ***	-0.459 ***	-0.083 **	0.532 ***	-0.588 ***	-0.016	0.598 ***	-0.345 ***	-0.021	0.678 ***	-0.542 ***	-0.034
	(0.194)	(0.072)	(0.025)	(0.087)	(0.045)	(0.037)	(0.105)	(0.051)	(0.011)	(0.130)	(0.051)	(0.030)	(0.084)	(0.040)	(0.023)
Separation Anxiety	-0.356 **	-0.051	-0.02												
	(0.142)	(0.043)	(0.014)												
Indirect Aggression				0.114	0.082 **	-0.03	-0.016	0.241 ***	0.015 *	0.288 ***	0.252 ***	-0.023	0.128	0.281 ***	0.088 ***
				0.081	0.033	0.024	0.101	0.041	0.009	0.098	0.042	0.025	0.138	0.043	0.021
Property Offences				0.396 ***	-0.291 ***	-0.218 ***				0.225 ***	-0.344 ***	0.004	0.185 ***	-0.306 ***	0.064 ***
				(0.083)	(0.050)	(0.035)				(0.079)	(0.033)	(0.020)	(0.071)	(0.026)	(0.013)

Note: All models are estimated using provided survey weights and bootstrap weights.
*p<0.1 , **p<0.05, ***p<0.01
Standard errors in parentheses.

Table B.4 Structural Model: Parameter Estimates Non-Cognitive Ability

PANEL A:	Latent Variables			
	Non-Cognitive Ability (θ_{t+1}^{NC})			
	C2 to C3	C3 to C4	C4 to C5	C5 to C6
Lagged Cognitive Ability (θ_t^C)	0.000	0.024	-0.042 **	-0.042 ***
	(0.027)	(0.018)	(0.018)	(0.016)
Lagged Non-Cognitive Ability (θ_t^N)	0.593 ***	0.764 ***	0.877 ***	0.857 ***
	(0.024)	(0.017)	(0.015)	(0.012)
Parental Inputs: Positive Interaction (I_t^1)	0.226 ***	0.187 ***	0.217 ***	0.214 ***
	(0.024)	(0.018)	(0.013)	(0.016)
Parental Inputs: Ineffective Parenting (I_t^2)	-0.539 ***	-0.549 ***	-0.475 ***	-0.514 ***
	(0.016)	(0.017)	(0.014)	(0.017)
Parental Inputs: Consistent Parenting (I_t^3)	0.277 ***	0.206 ***	0.276 ***	0.292 ***
	(0.019)	(0.016)	(0.017)	(0.014)

PANEL B:	Observed Covariates			
	Non-Cognitive Ability (θ_{t+1}^{NC})			
	C2 to C3	C3 to C4	C4 to C5	C5 to C6
Second Income Quintile	0.021	-0.011	0.050 **	0.113 ***
	(0.025)	(0.021)	(0.024)	(0.023)
Third Income Quintile	-0.069 **	-0.030	-0.032	0.032
	(0.035)	(0.027)	(0.025)	(0.027)
Fourth Income Quintile	0.018	-0.048	0.055 *	0.015
	(0.030)	(0.037)	(0.030)	(0.033)
Fifth Income Quintile	0.023	-0.019	0.061 *	0.132 ***
	(0.043)	(0.035)	(0.034)	(0.031)

Note: All models are estimated using provided survey weights and bootstrap weights.
*p<0.1 , **p<0.05, ***p<0.01
Standard errors in parentheses.

Table B.5 Structural Model: Parameter Estimates Cognitive Ability

PANEL A:	Latent Variables			
	Cognitive Ability (θ^C_{t+1})			
	C2 to C3	C3 to C4	C4 to C5	C5 to C6
Lagged Cognitive Ability (θ^C_t)	0.234 ***	0.421 ***	0.295 ***	0.482 ***
	(0.019)	(0.024)	(0.022)	(0.031)
Lagged Non-Cognitive Ability (θ^N_t)	0.033 **	0.077 ***	0.024	0.07 ***
	(0.013)	(0.014)	(0.016)	(0.015)
Parental Inputs: Positive Interaction (I^1_t)	0.102 ***	0.087 ***	0.015	0.017
	(0.014)	(0.014)	(0.013)	(0.013)
Parental Inputs: Ineffective Parenting (I^2_t)	-0.05 ***	-0.111 ***	0.001	-0.051 ***
	(0.012)	(0.016)	(0.010)	(0.011)
Parental Inputs: Consistent Parenting (I^3_t)	0.059 ***	0.045 ***	0.009	0.13 ***
	(0.015)	(0.016)	(0.013)	(0.013)

PANEL B:	Observed Covariates			
	Cognitive Ability (θ^C_{t+1})			
	C2 to C3	C3 to C4	C4 to C5	C5 to C6
Second Income Quintile	0.020	-0.095 ***	0.099 ***	0.048 ***
	(0.018)	(0.022)	(0.022)	(0.018)
Third Income Quintile	0.046 *	-0.204 ***	0.104 ***	-0.063 ***
	(0.024)	(0.027)	(0.025)	(0.022)
Fourth Income Quintile	0.120 ***	-0.191 ***	0.134 ***	-0.069 ***
	(0.025)	(0.026)	(0.026)	(0.021)
Fifth Income Quintile	0.079 ***	-0.183 ***	0.122 ***	-0.018
	(0.029)	(0.028)	(0.026)	(0.025)

Note: All models are estimated using provided survey weights and bootstrap weights.
*p<0.1 , **p<0.05, ***p<0.01
Standard errors in parentheses.